Stochastic maximum principle for optimal control of SPDEs

Marco Fuhrman
Dipartimento di Matematica, Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy
e-mail: marco.fuhrman@polimi.it

Ying Hu
IRMAR, Université Rennes 1
Campus de Beaulieu, 35042 Rennes Cedex, France
e-mail: ying.hu@univ-rennes1.fr

Gianmario Tessitore
Dipartimento di Matematica e Applicazioni, Università di Milano-Bicocca
Via Cozzi 53, 20125 Milano, Italy
e-mail: gianmario.tessitore@unimib.it

Abstract

We prove a version of the maximum principle, in the sense of Pontryagin, for the optimal control of a stochastic partial differential equation driven by a finite dimensional Wiener process. The equation is formulated in a semi-abstract form that allows direct applications to a large class of controlled stochastic parabolic equations. We allow for a diffusion coefficient dependent on the control parameter, and the space of control actions is general, so that in particular we need to introduce two adjoint processes. The second adjoint process takes values in a suitable space of operators on L^4.

1 Introduction

The problem of finding sufficient conditions for optimality for a stochastic optimal control problem with infinite dimensional state equation, along the lines of the Pontryagin maximum principle, was already addressed in the early 80's in the pioneering paper [1].

Despite the fact that the finite dimensional analogue of the problem has been solved, in complete generality, more than 20 years ago (see the well known paper by S. Peng [13]) the infinite dimensional case still has important open issues both on the side of the generality of the abstract model and on the side of its applicability to systems modeled by stochastic partial differential equations (SPDEs).

In particular, whereas the Pontryagin maximum principle for infinite dimensional stochastic control problems is a well known result as far as the control domain is convex (or the diffusion does not depend on the control), see [1, 8], for the general case (that is when the control domain need not be convex and the diffusion coefficient can contain a control variable) existing results are limited to abstract evolution equations under assumptions that are not satisfied by the large majority of concrete SPDEs.

The technical obstruction is related to the fact that (as it was pointed out in [13]) if the control domain is not convex the optimal control has to be perturbed by the so called "spike variation".

Then if the control enters the diffusion, the irregularity in time of the Brownian trajectories imposes to take into account a second variation process. Thus the stochastic maximum principle has to involve an adjoint process for the second variation. In the finite dimensional case such a process can be characterized as the solution of a matrix valued backward stochastic differential equation (BSDE) while in the infinite dimensional case the process naturally lives in a non-Hilbertian space of operators and its characterization is much more difficult. Moreover the applicability of the abstract results to concrete controlled SPDEs is another delicate step due to the specific difficulties that they involve such as the lack of regularity of Nemytskii-type coefficients in L^p spaces.

The present results (that were anticipated in the 6th International Symposium on BSDEs and Applications - Los Angeles, 2011 - and published in a short version in [5]) are, as far as we know, the only ones that can cover, for instance, a controlled stochastic heat equation (with finite dimensional noise) such as:

$$\begin{cases} dX_t(x) &= AX_t(x)\,dt + b(x, X_t(x), u_t)\,dt + \displaystyle\sum_{j=1}^{m} \sigma_j(x, X_t(x), u_t)\,d\beta_t^j, \qquad t \in [0,T], x \in \mathcal{O}, \\ X_0(x) &= x_0(x), \end{cases} \tag{1.1}$$

with $A = \Delta_x$ with appropriate boundary conditions, and a cost functional as follows:

$$J(u) = \mathbb{E}\int_0^T \int_{\mathcal{O}} l(x, X_t(x), u_t)\,dx\,dt + \mathbb{E}\int_{\mathcal{O}} h(x, X_T(x))\,dx,$$

$\mathcal{O} \subset \mathbb{R}^n$ being a bounded open set with regular boundary.

We stress the fact that in this paper the state equation is formulated, as above, only in a semi-abstract way in order, on one side, to cope with all the difficulties carried by the concrete non-linearities and on the other to take advantage of the regularizing properties of the leading elliptic operator.

Concerning other results on the infinite dimensional stochastic Pontryagin maximum principle, as we already mentioned in [1] and [8] the case of diffusion independent on the control is treated (with the difference that in [8] a complete characterization of the adjoint process to the first variation as the unique mild solution to a suitable BSDE is achieved). Then in [19] the case of linear state equation and cost functional is addressed. In this case as well, the second variation process is not needed.

The pioneering paper [14] is the first one in which the general case is addressed with, in addition, a general class of noises possibly with jumps. The adjoint process of the second variation $(P_t)_{t\in[0,T]}$ is characterized as the solution of a BSDE in the (Hilbertian) space of Hilbert Schmidt operators. This forces to assume a very strong regularity on the abstract state equation and control functional that prevents application of the general results to SPDEs. Recently in [10] P_t was characterized as "transposition solution" of a backward stochastic evolution equation in $\mathcal{L}(L^2(\mathcal{O}))$. Coefficients are required to be twice Fréchet-differentiable as operators in $L^2(\mathcal{O})$. Finally even more recently in a couple of preprints [3] [4] the process P_t is characterized in a similar way as it is in [5] and here. Roughly speaking it is characterized as a suitable stochastic bilinear form (see relation (5.6)). As it is the case in [10], in [3] and [4] as well the regularity assumptions on the coefficients are too restrictive to apply directly the general results to controlled SPDEs. On the other side in [4] an unbounded diffusion term is included in the model that can not be covered by the present results. Finally other variants of the problem have been studied. For instance in [6] a maximum principle for a SPDE with noise and control on the boundary but control independent diffusion is addressed, see also [11] for a case with delay.

The paper is structured as follows. In Section 2 we fix notations and standing assumptions. In Section 3 we state the main result. In Section 4 we recall the spike variation technique and introduce the first variation process, the corresponding adjoint process and the second variation process together with crucial estimates on them. We stress here the fact that the structure of the second variation process forces to develop a theory in the L^p spaces for the state equation and its perturbations through spike variation. In section 5 we complete the proof of the stochastic maximum principle. This is achieved by characterizing the adjoint of the second variation as a progressive process $(P_t)_{t\in[0,T]}$ with values in the space of linear bounded operators $L^4 \to (L^4)^* = L^{4/3}$. Namely P_t is defined through the stochastic bilinear form

$$\langle P_t f, g\rangle = \mathbb{E}^{\mathcal{F}_t} \int_t^T \langle \bar{H}_s Y_s^{t,f}, Y_s^{t,g}\rangle\, ds + \mathbb{E}^{\mathcal{F}_t} \langle \bar{h} Y_T^{t,f}, Y_T^{t,g}\rangle, \qquad \mathbb{P}-a.s.$$

where $(Y_s^{t,f})_{s\in[t,T]}$ is the mild solution of a suitable infinite dimensional forward stochastic equation (see equation (5.1)). The study of the regularity of process $(P_t)_{t\in[0,T]}$ is one of the main technical issues of this paper (together with the L^p estimates of the first and second variations) and exploits the specific properties of the semigroup generated by the elliptic differential operator. Finally in the Appendix we report some results on stochastic integration in L^p spaces. For the reader's convenience we give complete and direct proofs of some results (including a version of the Itô inequality, see (A.1)). Such results are particular cases of the ones obtained in the framework of stochastic calculus in UMD Banach spaces, see [17].

2 Notations and preliminaries

We begin by formulating an abstract form of the controlled PDE.

Let $(D, \mathcal{D}, m)$ be a measure space with finite measure (in the applications D is an open subset of $\mathbb{R}^N$ and m is the Lebesgue measure). We will consider the usual real spaces $L^p(D, \mathcal{D}, m)$, $p \in [1, \infty)$, which are shortly denoted by L^p and endowed with the usual norm $\|\cdot\|_p$.

Let $(W_t^1, \dots, W_t^d)_{t\geq 0}$ be a standard, d-dimensional Wiener process defined in some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $(\mathcal{F}_t)_{t\geq 0}$ the corresponding natural filtration, augmented in the usual way, and we denote by $\mathcal{P}$ the progressive σ-algebra on $\Omega \times [0, \infty)$ (or on a finite interval $[0, T]$, by abuse of notation). We will assume that there exist regular conditional probabilities $\mathbb{P}(\cdot|\mathcal{F}_t)$ given any $\mathcal{F}_t$: this holds for instance if the Wiener process is canonically realized on the space of $\mathbb{R}^d$-valued continuous functions.

As the space of control actions we take a separable metric space U, endowed with its Borel σ-algebra $\mathcal{B}(U)$. In general, we denote $\mathcal{B}(\Lambda)$ the Borel σ-algebra of any topological space Λ. We fix a finite time horizon $T > 0$ and by a control process we mean any progressive process $(u_t)_{t\in[0,T]}$ with values in U.

We consider the following controlled stochastic equation:

$$\begin{cases} dX_t(x) = AX_t(x)\, dt + b(t, x, X_t(x), u_t)\, dt + \sum_{j=1}^d \sigma_j(t, x, X_t(x), u_t)\, dW_t^j, \\ X_0(x) = x_0(x) \end{cases} \tag{2.1}$$

and the cost functional

$$J(u) = \mathbb{E} \int_0^T \int_D l(t, x, X_t(x), u_t)\, m(dx)\, dt + \mathbb{E} \int_D h(x, X_T(x))\, m(dx). \tag{2.2}$$

A control process u is called optimal if it minimizes the cost over all control processes. Denoting by X the corresponding trajectory we say that (u, X) is an optimal pair.

Hypothesis 2.1 *1. The operator A is the infinitesimal generator of a strongly continuous semigroup $(e^{tA})_{t\geq 0}$ of linear bounded operators on L^2. We assume that there exist constants $\bar{p} > 8$ and $M > 0$ such that for $p \in [2, \bar{p}]$ we have $e^{tA}(L^p) \subset L^p$ and $\|e^{tA}f\|_p \leq M\|f\|_p$ for every $t \in [0, T]$ and $f \in L^p$.*

2. *For $\phi = b$ or $\phi = l$ or $\phi = \sigma_j$, $j = 1, \dots, d$, the functions*

$$\phi(\omega, t, x, r, u) : \Omega \times [0, T] \times D \times \mathbb{R} \times U \to \mathbb{R}, \quad h(\omega, x, r) : \Omega \times D \times \mathbb{R} \to \mathbb{R},$$

are assumed to be measurable with respect to $\mathcal{P} \otimes \mathcal{D} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(U)$ and $\mathcal{B}(\mathbb{R})$ (respectively, $\mathcal{F}_T \otimes \mathcal{D} \otimes \mathcal{B}(\mathbb{R})$ and $\mathcal{B}(\mathbb{R})$).

3. *For every (ω, t, x, u), the functions $r \mapsto \phi(\omega, t, x, r, u)$ and $r \mapsto h(\omega, x, r)$ are continuous and have first and second derivatives, denoted ϕ' and ϕ'' (respectively, h' and h''), which are also continuous functions of r. We also assume that*

$$(|\phi'| + |\phi''| + |h'| + |h''|)(\omega, t, x, r, u) \leq K,$$

$$(|\phi| + |h|)(\omega, t, x, r, u) \leq K(|r| + |\bar{\psi}(x)|),$$

for some constant K, some $\bar{\psi} \in L^{\bar{p}}$ and for all (ω, t, x, r, u).

4. *$x_0 \in L^{\bar{p}}$.*

From now on we adopt the convention of summation over repeated indices, so that we will drop the symbol $\sum_{j=1}^{d}$ in (2.1).

Under the stated assumptions, for every control process u there exists a unique solution of the state equation (2.1) in the so-called mild sense, i.e. an adapted process $(X_t)_{t\in[0,T]}$ with values in L^2, with continuous trajectories, satisfying $\mathbb{P}$-a.s.

$$X_t = e^{tA}x_0 + \int_0^t e^{(t-s)A}b(s, \cdot, X_s(\cdot), u_s)\, ds + \int_0^t e^{(t-s)A}\sigma_j(s, \cdot, X_s(\cdot), u_s)\, dW_s^j, \qquad t \in [0, T]. \tag{2.3}$$

Here and below, equalities like (2.3) are understood to hold m-a.e., and uniqueness is understood up to modification of L^2-valued random processes.

We need to prove the following higher summability property of trajectories. We introduce the notation

$$|||X|||_p = \sup_{t\in[0,T]} (\mathbb{E}\|X_t\|_p^p)^{1/p}. \tag{2.4}$$

Proposition 2.2 *For every $p \in [2, \bar{p}]$, $(X_t)_{t\in[0,T]}$ is a progressive process with values in L^p, satisfying $\sup_{t\in[0,T]} \mathbb{E}\|X_t\|_p^p < \infty$.*

Proof. We consider the Banach space of progressive L^p-valued processes $(X_t)_{t\in[0,T]}$ such that the norm $|||X|||_p$ is finite. For such a process X we define

$$\Gamma(X)_t = e^{tA}x_0 + \int_0^t e^{(t-s)A}b(s, \cdot, X_s(\cdot), u_s)\, ds + \int_0^t e^{(t-s)A}\sigma_j(s, \cdot, X_s(\cdot), u_s)\, dW_s^j.$$

It can be proved that the map Γ is a contraction with respect to the norm $||| \cdot |||_p$, provided T is sufficiently small. Its unique fixed point is then the required solution. The restriction on T is then removed in a standard way by subdividing $[0, T]$ into appropriate subintervals.

The fact that Γ is a well defined contraction follows from moment estimates of the stochastic integrals in L^p. We limit ourselves to showing the contraction property assuming for simplicity $b = 0$. In this case, if $|||X|||_p + |||Y|||_p < \infty$, we have by (A.2)

$$\mathbb{E}\|\Gamma(X)_t - \Gamma(Y)_t\|_p^p \le c_p \int_0^t \mathbb{E}\|e^{(t-s)A}[\sigma(s,\cdot,X_s(\cdot),u_s) - \sigma(s,\cdot,Y_s(\cdot),u_s)]\|^p_{L^p(D;\mathbb{R}^d)} ds\ t^{(p-2)/2}.$$

Using the L^p-boundedness of e^{tA} and the Lipschitz character of σ which follows from Hypothesis 2.1-3 we obtain

$$\mathbb{E}\|\Gamma(X)_t - \Gamma(Y)_t\|_p^p \le C \int_0^t \mathbb{E}\|X_s - Y_s\|^p ds\ t^{(p-2)/2} \le C|||X - Y|||_p^p T^{p/2}$$

for some constant C independent of T. The contraction property follows immediately for T sufficiently small. □

3 Statement of the main result

3.1 Statement of the stochastic maximum principle

For our main result we also need the following assumptions.

Hypothesis 3.1 *There exists a complete orthonormal basis $(e_i)_{i\ge1}$ in L^2 which is also a Schauder basis of L^4.*

Hypothesis 3.2 *The restriction of $(e^{tA})_{t\ge0}$ to the space L^4 is a strongly continuous analytic semigroup and the domain of its infinitesimal generator is compactly embedded in L^4.*

We note that Hypothesis 3.1 is satisfied for a large class of measure spaces $(D, \mathcal{D}, m)$, typically with a basis of Haar type.

In the following a basic role will be played by the space of linear bounded operators $L^4 \to (L^4)^* = L^{4/3}$ endowed with the usual operator norm, that we simply denote by $\mathcal{L}$. Clearly, $\mathcal{L}$ may be identified with the space of bounded bilinear forms on L^4. The duality between $g \in L^4$ and $h \in L^{4/3}$ will be denoted $\langle h, g\rangle$. As it is customary when dealing with spaces of operators endowed with the operator norm, when considering random variables or processes with values in $\mathcal{L}$, the latter will be endowed with the Borel σ-algebra of the weak topology (the weakest topology making all the functions $T \mapsto \langle Tf, g\rangle$ continuous, $f, g \in L^4$); note that this is in general different from the Borel σ-algebra of the topology corresponding to the operator norm $\|T\|_{\mathcal{L}}$.

For $u \in U$ and $X, p, q^1, \dots, q^d \in L^2$ denote

$$\mathcal{H}(t,u,X,p,q^1,\dots,q^d) = \int_D [l(t,x,X(x),u) + b(t,x,X(x),u)p(x) + \sigma_j(t,x,X(x),u)q^j(x)]\, m(dx)$$

Theorem 3.3 *Let (X, u) be an optimal pair. Then there exist progressive processes $(P_t)_{t\in[0,T]}$ and $(p_t, q_t^1, \dots, q_t^d)_{t\in[0,T]}$, with values in $\mathcal{L}$ and $(L^2)^{d+1}$ respectively, for which the following inequality holds, $\mathbb{P}$-a.s. for a.e. $t \in [0,T]$: for every $v \in U$,*

$$\begin{aligned}&\mathcal{H}(t,v,X_t,p_t,q_t^1,\dots,q_t^d) - \mathcal{H}(t,u_t,X_t,p_t,q_t^1,\dots,q_t^d)\\ &\quad+\frac{1}{2}\langle P_t[\sigma_j(t,\cdot,X_t(\cdot),v) - \sigma_j(t,\cdot,X_t(\cdot),u_t)], \sigma_j(t,\cdot,X_t(\cdot),v) - \sigma_j(t,\cdot,X_t(\cdot),u_t)\rangle \ge 0.\end{aligned}$$

The process $(p, q^1, \dots, q^d)$ satisfies $\sup_{t\in[0,T]} \mathbb{E}\|p_t\|_2^2 + \mathbb{E}\int_0^T \sum_{j=1}^d \|q_t^j\|_2^2\, dt < \infty$, and it is the unique solution to equation (4.17) *below.*

The process P satisfies $\sup_{t\in[0,T]} \mathbb{E}\|P_t\|^2_{\mathcal{L}} < \infty$ and it is defined in Proposition 5.3 below (formula (5.6)*).*

$(p, q^1, \dots, q^d)$ and P will be called the first and second adjoint process, respectively.

3.2 Application to stochastic PDEs of parabolic type

The purpose of this short subsection is to show that the main result can be immediately applied to concrete cases of controlled stochastic PDE of parabolic type on domains of Euclidean space.

Let D be a bounded open subset of $\mathbb{R}^n$ with smooth boundary ∂D, and let m be the Lebesgue measure. Consider the following PDE of reaction-diffusion type

$$\begin{cases} dX_t(x) = \Delta X_t(x)\,dt + b(t,x,X_t(x),u_t)\,dt + \sum_{j=1}^{d} \sigma_j(t,x,X_t(x),u_t)\,dW_t^j, \quad t \in [0,T], x \in D, \\ X_t(x) = 0, \quad t \in [0,T], x \in \partial D, \\ X_0(x) = x_0(x), \quad x \in D, \end{cases} \tag{3.1}$$

and the cost functional (2.2). In this example the Wiener process, the space of control actions and the space of control processes are as before; on the coefficients b, σ_j, l, h we make the assumptions of Hypothesis 2.1, points 2 and 3; finally we suppose $x_0 \in L^{\bar{p}}$ for some $\bar{p} > 8$.

We claim that all the conclusions of Theorem 3.3 hold true.

Indeed, we can define the operator $A = \Delta$ as an unbounded operator in L^2 with domain $H^2(D) \cap H_0^1(D)$ (the standard Sobolev spaces). Then A generates a strongly continuous, analytic contraction semigroup in all the spaces L^p, $1 < p < \infty$, and the domain of A is compactly embedded: see e.g. [9] or [12]. Therefore Hypothesis 2.1, point 1, and Hypothesis 3.2 hold true. Finally, Hypothesis 3.1 is verified by a Haar basis.

By similar arguments instead of Δ one can consider more general operators of elliptic type with appropriate boundary conditions.

4 The spike variation method and the first adjoint process

Throughout this section we assume that Hypothesis 2.1 holds, whereas Hypotheses 3.1 and 3.2 will be needed only starting from the next section.

4.1 Spike variation method and expansion of the state and the cost

Suppose that u is an optimal control and X the corresponding optimal trajectory. We fix $t_0 \in (0,T)$ and $\epsilon > 0$ such that $[t_0, t_0+\epsilon] \subset (0,T)$, we fix a control process v and we introduce in the usual way the spike variation process

$$u_t^\epsilon = \begin{cases} v_t, & \text{if } t \in [t_0, t_0+\epsilon], \\ u_t, & \text{if } t \notin [t_0, t_0+\epsilon]. \end{cases}$$

We denote by X^ϵ the trajectory corresponding to u^ϵ. We are going to construct two L^2-valued stochastic processes, denoted Y^ϵ and Z^ϵ, in such a way that the difference $X - X^\epsilon - Y^\epsilon - Z^\epsilon$ is small (in the sense of Proposition 4.4 below) and the difference of the cost functional $J(u^\epsilon) - J(u)$ can be expressed in an appropriate form involving Y^ϵ and Z^ϵ up to a small remainder: see Proposition 4.5.

Define

$$\delta^\epsilon \sigma_j(t,x) \quad = \quad \sigma_j(t,x,X_t(x),u_t^\epsilon) - \sigma_j(t,x,X_t(x),u_t)$$

and consider the stochastic PDE

$$\begin{cases} dY_t^\epsilon(x) &= AY_t^\epsilon(x)\,dt + b'(t,x,X_t(x),u_t)Y_t^\epsilon(x)\,dt \\ & \quad +\sigma_j'(t,x,X_t(x),u_t)Y_t^\epsilon(x)\,dW_t^j + \delta^\epsilon\sigma_j(t,x)\,dW_t^j, \\ Y_0^\epsilon(x) &= 0. \end{cases} \tag{4.1}$$

By the standard theory of stochastic evolution equations in Hilbert spaces, see e.g. [2], there exists a unique solution to (4.1) in the mild sense, i.e. a progressive process $(Y_t^\epsilon)_{t\in[0,T]}$ with values in L^2, satisfying $\sup_{t\in[0,T]}\mathbb{E}\|Y_t^\epsilon\|_2^2<\infty$ and, for every $t\in[0,T]$,

$$\begin{aligned} Y_t^\epsilon \quad &= \quad \int_0^t e^{(t-s)A}[b'(s,\cdot,X_s(\cdot),u_s)Y_s^\epsilon(\cdot)+\delta^\epsilon b(s,\cdot)]\,ds \\ &\quad + \int_0^t e^{(t-s)A}[\sigma_j'(s,\cdot,X_s(\cdot),u_s)Y_s^\epsilon(\cdot)+\delta^\epsilon\sigma_j(s,\cdot)]\,dW_s^j, \qquad \mathbb{P}-a.s. \end{aligned} \tag{4.2}$$

For the sequel we need the following more precise result.

Proposition 4.1 *For every $p\in[2,\bar p]$, $(Y_t^\epsilon)_{t\in[0,T]}$ is a progressive process with values in L^p, satisfying*

$$|||Y^\epsilon|||_p=\sup_{t\in[0,T]}(\mathbb{E}\|Y_t^\epsilon\|_p^p)^{1/p}\le C\epsilon^{1/2}.$$

To prove this result we need the following lemma, that will be used several times.

Lemma 4.2 *Given $\mathcal{P}\otimes\mathcal{D}$-measurable processes $\bar a,\bar\alpha,\bar b^j,\bar\beta^j$ consider the linear equation:*

$$\left\{\begin{array}{lcl} dV_t(x) &=& AV_t(x)\,dt+\bar a(t,x)V_t(x)\,dt+\bar\alpha(t,x)\,dt+\bar b^j(t,x)V_t(x)\,dW_t^j+\bar\beta^j(t,x)\,dW_t^j, \\ V_0(x) &=& 0. \end{array}\right. \tag{4.3}$$

Suppose $\bar a,\bar b^j$ bounded and $p\in[2,\bar p]$. Then the following holds.

1. *There exists a unique solution to (4.3) in the mild sense, i.e. a progressive process $(V_t)_{t\in[0,T]}$ with values in L^p, satisfying*

$$|||V|||_p=\sup_{t\in[0,T]}(\mathbb{E}\|V_t\|_p^p)^{1/p}\le C\int_0^T(\mathbb{E}\|\bar\alpha_t\|_p^p)^{1/p}dt+C\left(\int_0^T(\mathbb{E}\|\bar\beta_t\|_p^p)^{2/p}\right)^{1/2} \tag{4.4}$$

   and, for every $t\in[0,T]$,

$$V_t=\int_0^t e^{(t-s)A}[\bar a(s,\cdot)V_s(\cdot)+\bar\alpha(s,\cdot)]\,ds+\int_0^t e^{(t-s)A}[\bar b_j(s,\cdot)V_s(\cdot)+\bar\beta^j(s,\cdot)]\,dW_s^j, \qquad \mathbb{P}-a.s.$$

   provided the right-hand side of (4.4) is finite.

2. *If in addition $\sup_{t\in[0,T]}\mathbb{E}(\|\bar\alpha_t\|_p^p+\|\bar\beta_t\|_p^p)<\infty$ and $\bar a,\bar b^j$ are supported in a time interval of length ϵ, then*

$$\sup_{t\in[0,T]}(\mathbb{E}\|V_t\|_p^p)^{1/p}\le C\,\epsilon\sup_{t\in[0,T]}(\mathbb{E}\|\bar\alpha_t\|_p^p)^{1/p}+C\sqrt{\epsilon}\sup_{t\in[0,T]}(\mathbb{E}\|\bar\beta_t\|_p^p)^{1/p}.$$

   or equivalently

$$|||V|||_p\le C\,\epsilon\,|||\bar\alpha|||_p+C\sqrt{\epsilon}\,|||\bar\beta|||_p. \tag{4.5}$$

3. *In the case $p=2$ we have*

$$\sup_{t\in[0,T]}\mathbb{E}\|V_t\|_2^2\le C\int_0^T\mathbb{E}\|\bar\alpha_t\|_2^2\,dt+C\int_0^T\mathbb{E}\|\bar\beta_t\|_2^2dt=C\,(\|\bar\alpha\|^2_{L^2(\Omega\times D\times[0,T])}+\|\bar\beta\|^2_{L^2(\Omega\times D\times[0,T])}). \tag{4.6}$$

In (4.4), (4.5), (4.6) we set $\|\bar\beta_t\|_p := \|\bar\beta_t\|_{L^p(D;\mathbb{R}^d)}$, *and the constant* C *depends on the bounds on* $\bar a, \bar b^j$, *on the semigroup* (e^{tA}), *and on* p *and* T.

Proof. We consider again the Banach space of progressive L^p-valued processes $(V_t)_{t\in[0,T]}$ endowed with the norm $|||V|||_p = \sup_{t\in[0,T]}(\mathbb{E}\|V_t\|_p^p)^{1/p}$. By the same arguments as in the proof of Proposition 2.2 we can prove that the map Γ defined as

$$\Gamma(V)_t = \int_0^t e^{(t-s)A}\bar a(s,\cdot)V_s(\cdot)\,ds + \int_0^t e^{(t-s)A}\bar b_j(s,\cdot)V_s(\cdot)\,dW_s^j$$

is a (linear) contraction with respect to the norm $|||\cdot|||_p$, provided T is sufficiently small. Therefore there exists a unique solution V and it satisfies the inequality

$$|||V|||_p \le C\,|||\int_0^\cdot e^{(\cdot-s)A}\bar\alpha(s)\,ds|||_p + |||\int_0^\cdot e^{(\cdot-s)A}\bar\beta^j(s)\,dW_s^j|||_p.$$

The inequality (4.4) follows from an estimate of those stochastic integrals, using (A.1) and the L^p-boundedness of e^{tA}. The restriction on T is then removed by subdividing $[0,T]$ into appropriate subintervals. Finally, (4.5) and (4.6) follow from (4.4) and the Hölder inequality. □

Proof of Proposition 4.1. This is an immediate corollary of the previous lemma, noting that $|||\delta^\epsilon\sigma_j|||_p \le C$ as a consequence of the linear growth condition on σ_j (Hypothesis 2.1-4) and the fact that $|||X|||_p < \infty$ by Proposition 2.2. □

Define

$$\begin{aligned}
\delta^\epsilon b(t,x) &= b(t,x,X_t(x),u_t^\epsilon) - b(t,x,X_t(x),u_t),\\
\delta^\epsilon b'(t,x) &= b'(t,x,X_t(x),u_t^\epsilon) - b'(t,x,X_t(x),u_t),\\
\delta^\epsilon \sigma_j'(t,x) &= \sigma_j'(t,x,X_t(x),u_t^\epsilon) - \sigma_j'(t,x,X_t(x),u_t),
\end{aligned}$$

and consider the following stochastic PDE:

$$\left\{\begin{aligned}
dZ_t^\epsilon(x) &= AZ_t^\epsilon(x)\,dt + b'(t,x,X_t(x),u_t)Z_t^\epsilon(x)\,dt\\
&\quad+\frac{1}{2}b''(t,x,X_t(x),u_t)Y_t^\epsilon(x)^2\,dt + \delta^\epsilon b(t,x)\,dt + \delta^\epsilon b'(t,x)Y_t^\epsilon(x)\,dt\\
&\quad+\sigma_j'(t,x,X_t(x),u_t)Z_t^\epsilon(x)\,dW_t^j + \frac{1}{2}\sigma_j''(t,x,X_t(x),u_t)Y_t^\epsilon(x)^2\,dW_t^j\\
&\quad+\delta^\epsilon\sigma_j'(t,x)Y_t^\epsilon(x)\,dW_t^j,\\
Z_0^\epsilon(x) &= 0
\end{aligned}\right. \tag{4.7}$$

By the standard theory there exists a unique solution to (4.7) in the mild sense, i.e. a progressive process $(Z_t^\epsilon)_{t\in[0,T]}$ with values in L^2, satisfying $\sup_{t\in[0,T]}\mathbb{E}\|Z_t^\epsilon\|_2^2 < \infty$ and, for every $t\in[0,T]$,

$$\begin{aligned}
Z_t^\epsilon &= \int_0^t e^{(t-s)A}[b'(s,\cdot,X_s(\cdot),u_s)Z_s^\epsilon(\cdot) + \frac{1}{2}b''(s,\cdot,X_s(\cdot),u_s)Y_s^\epsilon(\cdot)^2 + \delta^\epsilon b(s,\cdot) + \delta^\epsilon b'(s,\cdot)Y_s^\epsilon(\cdot)]\,ds\\
&\quad+\int_0^t e^{(t-s)A}[\sigma_j'(s,\cdot,X_s(\cdot),u_s)Z_s^\epsilon(\cdot) + \frac{1}{2}\sigma_j''(s,\cdot,X_s(\cdot),u_s)Y_s^\epsilon(\cdot)^2 + \delta^\epsilon\sigma_j'(s,\cdot)Y_s^\epsilon(\cdot)]\,dW_s^j,\ \mathbb{P}\text{-}a.s.
\end{aligned} \tag{4.8}$$

For the sequel we need the following result.

Proposition 4.3 *For every* $p\in[2,\bar p/2]$, $(Z_t^\epsilon)_{t\in[0,T]}$ *is a progressive process with values in* L^p, *satisfying*

$$|||Z^\epsilon|||_p = \sup_{t\in[0,T]}(\mathbb{E}\|Z_t^\epsilon\|_p^p)^{1/p} \le C\epsilon.$$

Proof. The result follows from Lemma 4.2 applied to equation (4.7). In particular inequality (4.4) shows that

$$\begin{aligned} |||Z^\epsilon|||_p &\leq C\int_0^T (\mathbb{E}\|\frac{1}{2}b''(t,X_t,u_t)(Y_t^\epsilon)^2 + \delta^\epsilon b(t) + \delta^\epsilon b'(t)Y_t^\epsilon\|_p^p)^{1/p} dt \\ &\quad + C\left(\int_0^T (\mathbb{E}\|\frac{1}{2}\sigma''(t,X_t,u_t)(Y_t^\epsilon)^2 + \delta^\epsilon \sigma'(t)Y_t^\epsilon\|_p^p)^{2/p}\right)^{1/2}. \end{aligned}$$

The proof is now concluded estimating the right-hand side of this inequality.

Since $\|\sigma''(t,X_t,u_t)(Y_t^\epsilon)^2\|_p \leq C\|(Y_t^\epsilon)^2\|_p = C\|Y_t^\epsilon\|_{2p}^2$ we have

$$\left(\int_0^T (\mathbb{E}\|\sigma''(t,X_t,u_t)(Y_t^\epsilon)^2\|_p^p)^{2/p}\right)^{1/2} \leq C\left(\int_0^T (\mathbb{E}\|Y_t^\epsilon\|_{2p}^{2p})^{2/p}\right)^{1/2} \leq C|||Y^\epsilon|||_{2p}^2 \leq C\epsilon,$$

by Proposition 4.1, since $2p \leq \bar{p}$.

Next we note that $|||\delta^\epsilon b|||_p \leq C$, as a consequence of the linear growth condition on b (Hypothesis 2.1-3) and the fact that $|||X|||_p < \infty$ by Proposition 2.2. Since $\delta^\epsilon b$ is supported in $[t_0, t_0+\epsilon]$ it follows that $\int_0^T (\mathbb{E}\|\delta^\epsilon b(t)\|_p^p)^{1/p} dt \leq C\epsilon$.

The other terms are treated in a similar way. □

Proposition 4.4 *We have*

$$\sup_{t\in[0,T]} (\mathbb{E}\|X_t^\epsilon - X_t - Y_t^\epsilon - Z_t^\epsilon\|_2^2)^{1/2} = o(\epsilon).$$

As usual, $o(\epsilon)$ denotes any function of ϵ such that $o(\epsilon)/\epsilon \to 0$ as $\epsilon \to 0$. During the proof we will use the Taylor formula in the following form: for a twice continuously differentiable real function g on $\mathbb{R}$, and for $r, h \in \mathbb{R}$,

$$g(r+h) = g(r) + g'(r)h + \int_0^1\int_0^1 g''(r+\lambda\mu h)\,\mu d\lambda d\mu\, h^2. \tag{4.9}$$

Since $\int_0^1\int_0^1 \mu d\lambda d\mu = 1/2$ this can also be written

$$g(r+h) = g(r) + g'(r)h + \frac{1}{2}g''(r)h^2 + \int_0^1\int_0^1 [g''(r+\lambda\mu h) - g''(r)]\,\mu d\lambda d\mu\, h^2. \tag{4.10}$$

Proof. We set $R^\epsilon = Y^\epsilon + Z^\epsilon$. We first show that $X + R^\epsilon$ is a solution of the following equation in L^2:

$$\begin{aligned} X_t + R_t^\epsilon &= e^{tA}x_0 + \int_0^t e^{(t-s)A} b(s,\cdot,X_s(\cdot)+R_s^\epsilon(\cdot),u_s)\,ds - \int_0^t e^{(t-s)A}G^\epsilon(s,\cdot)\,ds \\ &+ \int_0^t e^{(t-s)A}\sigma_j(s,\cdot,X_s(\cdot)+R_s^\epsilon(\cdot),u_s)\,dW_s^j - \int_0^t e^{(t-s)A}\Lambda_j^\epsilon(s,\cdot)\,dW_s^j, \end{aligned} \tag{4.11}$$

where $G^\epsilon = G^{\epsilon,1} + G^{\epsilon,2} + G^{\epsilon,3}$, $\Lambda_j^\epsilon = \Lambda_j^{\epsilon,1} + \Lambda_j^{\epsilon,2} + \Lambda_j^{\epsilon,3}$,

$$\begin{aligned} G^{\epsilon,1}(s,x) &= \int_0^1\int_0^1 [b''(s,x,X_s(x)+\lambda\mu R_s^\epsilon(x),u_s^\epsilon) - b''(s,x,X_s(x),u_s)]\,\mu d\lambda d\mu\, R_s^\epsilon(x)^2, \\ G^{\epsilon,2}(s,x) &= \frac{1}{2}b''(s,x,X_s(x),u_s)\,(Z_s^\epsilon(x)^2 + 2Y_s^\epsilon(x)Z_s^\epsilon(x)), \qquad G^{\epsilon,3}(s,x) = \delta^\epsilon b'(s,x)Z_s^\epsilon(x), \\ \Lambda_j^{\epsilon,1}(s,x) &= \int_0^1\int_0^1 [\sigma_j''(s,x,X_s(x)+\lambda\mu R_s^\epsilon(x),u_s^\epsilon) - \sigma_j''(s,x,X_s(x),u_s)]\,\mu d\lambda d\mu\, R_s^\epsilon(x)^2, \\ \Lambda_j^{\epsilon,2}(s,x) &= \frac{1}{2}\sigma_j''(s,x,X_s(x),u_s)\,(Z_s^\epsilon(x)^2 + 2Y_s^\epsilon(x)Z_s^\epsilon(x)), \qquad \Lambda_j^{\epsilon,3}(s,x) = \delta^\epsilon\sigma_j'(s,x)Z_s^\epsilon(x). \end{aligned}$$

To verify (4.11) we use the Taylor formula (4.9) and obtain

$$\begin{aligned} b(s,x,X_s(x)+R_s^\epsilon(x),u_s^\epsilon) &= b(s,x,X_s(x),u_s^\epsilon)+b'(s,x,X_s(x),u_s^\epsilon)R_s^\epsilon(x) \\ &\quad +\int_0^1\int_0^1 b''(s,x,X_s(x)+\lambda\mu R_s^\epsilon(x),u_s^\epsilon)\,\mu d\lambda d\mu\, R_s^\epsilon(x)^2, \quad (4.12) \\ \sigma_j(s,x,X_s(x)+R_s^\epsilon(x),u_s^\epsilon) &= \sigma_j(s,x,X_s(x),u_s^\epsilon)+\sigma_j'(s,x,X_s(x),u_s^\epsilon)R_s^\epsilon(x) \\ &\quad +\int_0^1\int_0^1 \sigma_j''(s,x,X_s(x)+\lambda\mu R_s^\epsilon(x),u_s^\epsilon)\,\mu d\lambda d\mu\, R_s^\epsilon(x)^2. \quad (4.13) \end{aligned}$$

We apply $e^{(t-s)A}$ to (4.12) and integrate $\int_0^t ds$, we apply $e^{(t-s)A}$ to (4.13) and integrate $\int_0^t dW_s^j$, and we add the resulting equalities. Comparing with (2.3), (4.2) and (4.8) we obtain (4.11).

Since X^ϵ is the trajectory corresponding to u^ϵ we have

$$X_t^\epsilon = e^{tA}x_0+\int_0^t e^{(t-s)A}b(s,\cdot,X_s^\epsilon(\cdot),u_s^\epsilon)\,ds+\int_0^t e^{(t-s)A}\sigma_j(s,\cdot,X_s^\epsilon(\cdot),u_s^\epsilon)\,dW_s^j.$$

Comparing with (4.11) we see that $\Delta^\epsilon := X^\epsilon - X - R^\epsilon$ solves

$$\begin{aligned} \Delta_t^\epsilon = &\int_0^t e^{(t-s)A}\bar b^\epsilon(s,\cdot)\Delta_s^\epsilon(\cdot)\,ds+\int_0^t e^{(t-s)A}G^\epsilon(s,\cdot)\,ds \\ &+\int_0^t e^{(t-s)A}\bar\sigma_j^\epsilon(s,\cdot)\Delta_s^\epsilon(\cdot)\,dW_s^j+\int_0^t e^{(t-s)A}\Lambda_j^\epsilon(s,\cdot)\,dW_s^j, \end{aligned}$$

where

$$\begin{aligned} \bar b^\epsilon(s,x) &= \int_0^1 b'(s,x,X_s(x)+R_s^\epsilon(x)+\lambda\Delta_s^\epsilon(x),u_s^\epsilon)\,d\lambda, \\ \bar\sigma_j^\epsilon(s,x) &= \int_0^1 \sigma_j'(s,x,X_s(x)+R_s^\epsilon(x)+\lambda\Delta_s^\epsilon(x),u_s^\epsilon)\,d\lambda, \end{aligned}$$

are bounded coefficients, uniformly in ϵ. We can then apply Lemma 4.2 and specifically inequality (4.6) arriving at

$$\sup_{t\in[0,T]}(\mathbb{E}\|X_t^\epsilon - X_t - Y_t^\epsilon - Z_t^\epsilon\|_2^2)^{1/2} = \sup_{t\in[0,T]}\mathbb{E}\|\Delta_t^\epsilon\|_2^2 \le C\,(\|G^\epsilon\|^2_{L^2(\Omega\times D\times[0,T])}+\|\Lambda^\epsilon\|^2_{L^2(\Omega\times D\times[0,T])}).$$

To finish the proof it remains to verify that the $L^2(\Omega\times D\times[0,T])$-norm of each term $G^{\epsilon,i}$, $\Lambda_j^{\epsilon,i}$ $(i=1,2,3)$ is $o(\epsilon)$.

Let us verify that $\|G^{\epsilon,1}\|_{L^2(\Omega\times D\times[0,T])}=o(\epsilon)$. Write $G^{\epsilon,1}(t,x)=Q_t^\epsilon(x)R_t^\epsilon(x)^2$ where

$$Q_t^\epsilon(x)=\int_0^1\int_0^1 [b''(t,x,X_t(x)+\lambda\mu R_t^\epsilon(x),u_t^\epsilon)-b''(t,x,X_t(x),u_t)]\,\mu d\lambda d\mu.$$

Next take $p\in(2,\bar p/4]$, which is possible because $\bar p>8$, and let $q>1$ be such that $\frac{1}{2}=\frac{1}{p}+\frac{1}{q}$. Then

$$\|G^{\epsilon,1}\|_{L^2(\Omega\times D\times[0,T])}\le \|Q^\epsilon\|_{L^q(\Omega\times D\times[0,T])}\|(R^\epsilon)^2\|_{L^p(\Omega\times D\times[0,T])}.$$

Since

$$\|(R^\epsilon)^2\|_{L^p(\Omega\times D\times[0,T])}=\|R^\epsilon\|^2_{L^{2p}(\Omega\times D\times[0,T])}\le C|||R^\epsilon|||_{2p}^2\le C(|||Y^\epsilon|||_{2p}^2+|||Z^\epsilon|||_{2p}^2)\le C(\epsilon+\epsilon^2)$$

by Propositions 4.1 and 4.3, it remains to show that $\|Q^\epsilon\|_{L^q(\Omega\times D\times[0,T])}\to 0$ as $\epsilon\to 0$. We argue by contradiction: assume that there exists $\delta>0$ and a sequence $\epsilon_n\to 0$ such that

$$\mathbb{E}\int_0^T\int_D\left|\int_0^1\int_0^1[b''(t,x,X_t(x)+\lambda\mu R_t^\epsilon(x),u_t^\epsilon)-b''(t,x,X_t(x),u_t)]\,\mu d\lambda d\mu\right|^q m(dx)dt\ge\delta. \quad (4.14)$$

Since $\|R^\epsilon\|_{L^2(\Omega\times D\times[0,T])} \le C|||R^\epsilon|||_2 \le C(|||Y^\epsilon|||_2 + |||Z^\epsilon|||_2) \to 0$, there exists a subsequence ϵ_{n_k} such that $R^{\epsilon_{n_k}} \to 0$ a.s. with respect to the product measure $\mathbb{P}(d\omega)m(dx)dt$. Since $r \mapsto b''(t,x,r,u)$ is continuous, and due to the special definition of u^ϵ, it follows that $b''(t,x,X_t(x) + \lambda\mu R_t^{\epsilon_{n_k}}(x), u_t^{\epsilon_{n_k}}) \to b''(t,x,X_t(x),u_t)$ a.s. with respect to

$\mathbb{P}(d\omega)m(dx)dtd\lambda d\mu$. By dominated convergence this contradicts (4.14).

The proof that $\|\Lambda_j^{\epsilon,1}\|_{L^2(\Omega\times D\times[0,T])} \to 0$ is identical. The other terms $G^{\epsilon,i}$, $\Lambda_j^{\epsilon,i}$ are treated in a standard way using Propositions 4.1 and 4.3. □

Define

$$\begin{aligned}
\delta^\epsilon l(t,x) &= l(t,x,X_t(x),u_t^\epsilon) - l(t,x,X_t(x),u_t), \\
\delta^\epsilon l'(t,x) &= l'(t,x,X_t(x),u_t^\epsilon) - l'(t,x,X_t(x),u_t).
\end{aligned}$$

Proposition 4.5 *We have*

$$\begin{aligned}
&J(u^\epsilon) - J(u) = \mathbb{E}\int_0^T \int_D \delta^\epsilon l(t,x)\, m(dx)\, dt \\
&+\mathbb{E}\int_0^T \int_D l'(t,x,X_t(x),u_t)(Y_t^\epsilon(x) + Z_t^\epsilon(x))\, m(dx)\, dt \\
&+\frac{1}{2}\mathbb{E}\int_0^T \int_D l''(t,x,X_t(x),u_t)Y_t^\epsilon(x)^2\, m(dx)\, dt \\
&+\mathbb{E}\int_D h'(x,X_t(x))(Y_T^\epsilon(x) + Z_T^\epsilon(x))\, m(dx) + \frac{1}{2}\mathbb{E}\int_D h''(x,X_T(x))Y_T^\epsilon(x)^2\, m(dx) + o(\epsilon).
\end{aligned} \tag{4.15}$$

Proof. We still denote $R^\epsilon = Y^\epsilon + Z^\epsilon$. We have

$$\begin{aligned}
J(u^\epsilon) - J(u) = {}& \mathbb{E}\int_0^T \int_D [l(t,x,X_t^\epsilon(x),u_t^\epsilon) - l(t,x,X_t(x),u_t)]\, m(dx)\, dt \\
&+\mathbb{E}\int_D [h(x,X_T^\epsilon(x)) - h(x,X_T(x))]\, m(dx).
\end{aligned}$$

We first consider

$$\mathbb{E}\int_0^T \int_D [l(t,x,X_t^\epsilon(x),u_t^\epsilon) - l(t,x,X_t(x),u_t)]\, m(dx)\, dt = A_1 + A_2 + A_3,$$

where

$$\begin{aligned}
A_1 &= \mathbb{E}\int_0^T \int_D [l(t,x,X_t^\epsilon(x),u_t^\epsilon) - l(t,x,X_t(x) + R_t^\epsilon(x),u_t^\epsilon)]\, m(dx)\, dt, \\
A_2 &= \mathbb{E}\int_0^T \int_D [l(t,x,X_t(x) + R_t^\epsilon(x),u_t^\epsilon) - l(t,x,X_t(x) + R_t^\epsilon(x),u_t)]\, m(dx)\, dt, \\
A_3 &= \mathbb{E}\int_0^T \int_D [l(t,x,X_t(x) + R_t^\epsilon(x),u_t) - l(t,x,X_t(x),u_t)]\, m(dx)\, dt.
\end{aligned}$$

From Proposition 4.4 it follows that $A_1 = o(\epsilon)$. Next applying the Taylor formula (4.9) twice in A_2 we have

$$\begin{aligned}
&A_2 = \mathbb{E}\int_0^T \int_D \Bigg(\delta^\epsilon l(t,x) + \delta^\epsilon l'(t,x)R_t^\epsilon(x) \\
&+\int_0^1 \int_0^1 [l''(t,x,X_t(x) + \lambda\mu R_t^\epsilon(x),u_t^\epsilon) - l''(t,x,X_t(x) + \lambda\mu R_t^\epsilon(x),u_t)]\, \mu d\lambda d\mu\, R_t^\epsilon(x)^2\Bigg)\, m(dx)\, dt \\
&= \mathbb{E}\int_0^T \int_D \delta^\epsilon l(t,x)\, m(dx)\, dt + o(\epsilon),
\end{aligned}$$

as it follows easily from Propositions 4.3 and 4.3. Applying the Taylor formula (4.10) we have

$$\begin{aligned}
A_3 &= \mathbb{E}\int_0^T \int_D \Big(l'(t,x,X_t(x),u_t)R_t^\epsilon(x) + \frac{1}{2} l''(t,x,X_t(x),u_t)R_t^\epsilon(x)^2 \\
&+ \int_0^1\int_0^1 [l''(t,x,X_t(x)+\lambda\mu R_t^\epsilon(x),u_t) - l''(t,x,X_t(x),u_t)]\,\mu d\lambda d\mu\, R_t^\epsilon(x)^2 \Big)\, m(dx)\,dt \\
&= \mathbb{E}\int_0^T \int_D \Big(l'(t,x,X_t(x),u_t)R_t^\epsilon(x) + \frac{1}{2} l''(t,x,X_t(x),u_t)Y_t^\epsilon(x)^2 \Big)\, m(dx)\,dt + o(\epsilon).
\end{aligned}$$

The last equality is verified noting that

$$\mathbb{E}\int_0^T \int_D l''(t,x,X_t(x),u_t)\,(2Y_t^\epsilon(x)Z_t^\epsilon(x) + Z_t^\epsilon(x)^2)\, m(dx)\,dt = o(\epsilon),$$

by Propositions 4.3 and 4.3, and that

$$\mathbb{E}\int_0^T\int_D\Big(\int_0^1\int_0^1 [l''(t,x,X_t(x)+\lambda\mu R_t^\epsilon(x),u_t) - l''(t,x,X_t(x),u_t)]\,\mu d\lambda d\mu\, R_t^\epsilon(x)^2\Big)\, m(dx)\,dt = o(\epsilon)$$

which can be proved by the same arguments used to treat the term $G^{\epsilon,1}$ in the proof of Proposition 4.4.

In a similar way one proves

$$\begin{aligned}
&\mathbb{E}\int_D [h(x,X_T^\epsilon(x)) - h(x,X_T(x))]\, m(dx) \\
&= \mathbb{E}\int_D h'(x,X_t(x))(Y_T^\epsilon(x)+Z_T^\epsilon(x))\, m(dx) + \frac{1}{2}\mathbb{E}\int_D h''(x,X_T(x))Y_T^\epsilon(x)^2\, m(dx) + o(\epsilon),
\end{aligned}$$

and the proof is finished. □

4.2 The first adjoint process

The first adjoint process is defined as the solution of the backward stochastic PDE

$$\left\{\begin{array}{rcl}
-dp_t(x) &=& -dq_t^j(x)\,dW_t^j + [A^*p_t(x) + b'(t,x,X_t(x),u_t)p_t(x) \\
&& +\sigma_j'(t,x,X_t(x),u_t)q_t^j(x) + l'(t,x,X_t(x),u_t)]\,dt \\
p_T(x) &=& h'(x,X_T(x))
\end{array}\right. \tag{4.16}$$

where A^* denotes the adjoint of A in L^2. By the result in [7] there exists a unique solution, i.e. a progressive process $(p_t, q_t^1, \dots, q_t^d)_{t\in[0,T]}$ with values in $(L^2)^{d+1}$, such that

$$\sup_{t\in[0,T]} \mathbb{E}\|p_t\|_2^2 + \mathbb{E}\int_0^T \sum_{j=1}^d \|q_t^j\|_2^2\,dt < \infty,$$

and satisfying the equation in the mild sense: for every $t\in[0,T]$,

$$\begin{aligned}
p_t + \int_t^T e^{(s-t)A^*} q_s^j\,dW_s^j &= e^{(T-t)A^*}h'(\cdot,X_T(\cdot)) + \int_t^T e^{(s-t)A^*}[b'(s,\cdot,X_s(\cdot),u_s)p_s(\cdot) \\
&\quad +\sigma_j'(s,\cdot,X_s(\cdot),u_s)q_s^j(\cdot) + l'(s,\cdot,X_s(\cdot),u_s)]\,ds,\ \mathbb{P}-a.s.
\end{aligned} \tag{4.17}$$

where (e^{tA^*}) denotes the adjoint semigroup of (e^{tA}) in L^2, which admits A^* as its generator.

Proposition 4.6 *Define*

$$\begin{array}{rcl} \bar{H}(t,x) & = & l''(t,x,X_t(x),u_t) + p_t(x) b''(t,x,X_t(x),u_t) + q_t^j(x)\sigma_j''(t,x,X_t(x),u_t), \\ \bar{h}(x) & = & h''(x,X_T(x)). \end{array} \tag{4.18}$$

Then we have

$$\begin{array}{rcl} J(u^\epsilon) - J(u) & = & \mathbb{E}\displaystyle\int_0^T \int_D [\delta^\epsilon l(t,x) + p_t(x)\delta^\epsilon b(t,x) + q_t^j(x)\delta^\epsilon \sigma_j(t,x)]\, ds\, m(dx) \\ & & +\dfrac{1}{2}\mathbb{E}\displaystyle\int_0^T \int_D \bar{H}(t,x)\, Y_t^\epsilon(x)^2\, ds\, m(dx) + \dfrac{1}{2}\mathbb{E}\int_D \bar{h}(x)\, Y_T^\epsilon(x)^2\, m(dx) + o(\epsilon). \end{array}$$

Proof. We claim that the following duality relations hold:

$$\begin{array}{l} \mathbb{E}\displaystyle\int_0^T \int_D l'(t,x,X_t(x),u_t) Y_t^\epsilon(x)\, m(dx)\, dt + \mathbb{E}\int_D h'(x,X_t(x)) Y_T^\epsilon(x)\, m(dx) \\ \quad = \mathbb{E}\displaystyle\int_0^T \int_D \delta^\epsilon \sigma_j(t,x) q_t^j(x)\, m(dx)\, dt, \end{array} \tag{4.19}$$

$$\begin{array}{l} \mathbb{E}\displaystyle\int_0^T \int_D l'(t,x,X_t(x),u_t) Z_t^\epsilon(x)\, m(dx)\, dt + \mathbb{E}\int_D h'(x,X_t(x)) Z_T^\epsilon(x)\, m(dx) \\ \quad = \mathbb{E}\displaystyle\int_0^T \int_D [\delta^\epsilon b(t,x) + \frac{1}{2} b''(t,x,X_t(x),u_t) Y_t^\epsilon(x)^2 + \delta^\epsilon b'(t,x) Y_t^\epsilon(x)] p_t(x)\, m(dx)\, dt \\ \quad +\mathbb{E}\displaystyle\int_0^T \int_D [\frac{1}{2}\sigma_j''(t,x,X_t(x),u_t) Y_t^\epsilon(x)^2 + \delta^\epsilon \sigma_j'(t,x) Y_t^\epsilon(x)] q_t^j(x)\, m(dx)\, dt. \end{array} \tag{4.20}$$

If A is a bounded operator and equations (4.1) and (4.7) are valid in the sense of Ito differentials in L^2 then (4.19) and (4.20) follow from an application of the Ito formula to the processes $\langle Y_t^\epsilon, p_t\rangle_{L^2}$ and $\langle Z_t^\epsilon, p_t\rangle_{L^2}$ respectively, where $\langle \cdot,\cdot\rangle_{L^2}$ denotes the scalar product in L^2. In the general case a regularization procedure is needed, where in particular the operator A is replaced by its Yosida approximation A_n and then $n \to \infty$. We omit writing down this standard part of the proof: one can find the details of these arguments (applied to BSDEs) in [16] or (applied to forward SDEs and control problems) in [15]. One can also look at Subsection 5.2 below where we use similar arguments in a more complicated setting.

Now the proof is concluded substituting (4.19) and (4.20) in (4.15), provided we can prove

$$\mathbb{E}\int_0^T \int_D \delta^\epsilon b'(t,x) Y_t^\epsilon(x) p_t(x)\, m(dx)\, dt = o(\epsilon), \quad \mathbb{E}\int_0^T \int_D \delta^\epsilon \sigma_j'(t,x) Y_t^\epsilon(x) q_t^j(x)\, m(dx)\, dt = o(\epsilon).$$

Since the proof is very similar, we only prove the second equality. Since $\delta^\epsilon \sigma_j'$ is bounded and supported in $[t_0, t_0+\epsilon]$ we have, using the Hölder inequality and recalling the norm $|||\cdot|||_p$ introduced in (2.4),

$$\begin{array}{l} \left|\mathbb{E}\displaystyle\int_0^T \int_D \delta^\epsilon \sigma_j'(t,x) Y_t^\epsilon(x) q_t^j(x)\, m(dx)\, dt\right| \le C\mathbb{E}\int_0^T 1_{[t_0,t_0+\epsilon]}(t) \|Y_t^\epsilon\|_2 \|q_t\|_2\, dt \\ \quad \le C|||Y^\epsilon|||_2 \displaystyle\int_0^T 1_{[t_0,t_0+\epsilon]}(t) (\mathbb{E}\|q_t\|_2^2)^{1/2}\, dt \le C\, |||Y^\epsilon|||_2 \left(\int_0^T 1_{[t_0,t_0+\epsilon]}(t)\mathbb{E}\|q_t\|_2^2\, dt\right)^{1/2} \sqrt{\epsilon}. \end{array}$$

The last integral tends to 0 as $\epsilon \to 0$, since $\mathbb{E}\int_0^T \|q_t\|_2^2\, dt < \infty$. It follows that the right-hand side is $o(\epsilon)$, because $|||Y^\epsilon|||_2 \le C\sqrt{\epsilon}$ by Proposition 4.1. $\square$

4.3 Some formal computations and heuristics

In order to motivate some of the constructions below, and to make a connection with the finite-dimensional case treated in [13], in this paragraph we proceed in a formal way.

We wish to prove that

$$\begin{aligned} &\mathbb{E}\int_0^T\int_D \bar{H}(t,x)Y_t^\epsilon(x)^2\,ds\,m(dx)+\mathbb{E}\int_D \bar{h}(x)Y_T^\epsilon(x)^2\,m(dx)\\ &=\mathbb{E}\int_0^T\langle P_t\delta^\epsilon\sigma_j(t,\cdot),\delta^\epsilon\sigma_j(t,\cdot)\rangle_{L^2}\,dt+o(\epsilon),\end{aligned}\tag{4.21}$$

for an appropriate operator-valued process P_t. In view of Proposition 4.6 the stochastic maximum principle then can be shown to hold by the usual arguments as in [13] or [18].

We denote by H_t the multiplication operator by the function $\bar{H}(t,\cdot)$ and by h the multiplication operator by the function $\bar{h}(\cdot)$. We pretend that they are bounded operators on the space L^2.

Next we consider the operator-valued BSDE

$$\left\{\begin{array}{lcl} -dP_t &=& -Q_t^j\,dW_t^j+[A^*P_t+P_tA+B_tP_t+P_tB_t+C_t^jP_tC_t^j+C_t^jQ_t^j+Q_t^jC_t^j+H_t]\,dt\\ P_T &=& h,\end{array}\right.\tag{4.22}$$

where by B_t, C_t^j we denote the (self-adjoint) multiplication operators by $b'(t,\cdot,X_t(\cdot),u_t)$ and $\sigma_j'(t,\cdot,X_t(\cdot),u_t)$ respectively. Suppose that we can find a good solution in the space of bounded linear operators on L^2. Then applying the Ito formula to $\langle P_tY_t^\epsilon,Y_t^\epsilon\rangle_{L^2}$, integrating from 0 to T and taking expectations we obtain

$$\begin{aligned} &\mathbb{E}\int_0^T\langle H_tY_t^\epsilon,Y_t^\epsilon\rangle_{L^2}\,dt+\mathbb{E}\langle hY_T^\epsilon,Y_T^\epsilon\rangle_{L^2}=\mathbb{E}\int_0^T[\langle P_t\delta^\epsilon\sigma_j(t,\cdot),\delta^\epsilon\sigma_j(t,\cdot)\rangle_{L^2}\\ &+2\langle P_t\delta^\epsilon b(t,\cdot),Y_t^\epsilon\rangle_{L^2}+2\langle P_tC_t^jY_t^\epsilon,\delta^\epsilon b(t,\cdot)\rangle_{L^2}-2\langle Q_t^jY_t^\epsilon,\delta^\epsilon\sigma_j(t,\cdot)\rangle_{L^2}]\,dt.\end{aligned}$$

If we were able to prove, in analogy with the finite-dimensional case, that

$$\mathbb{E}\int_0^T[2\langle P_t\delta^\epsilon b(t,\cdot),Y_t^\epsilon\rangle_{L^2}+2\langle P_tC_t^jY_t^\epsilon,\delta^\epsilon b(t,\cdot)\rangle_{L^2}-2\langle Q_t^jY_t^\epsilon,\delta^\epsilon\sigma_j(t,\cdot)\rangle_{L^2}]\,dt=o(\epsilon),$$

then (4.21) would follow and the proof would be finished. However in this argument finding a solution of the operator-valued BSDE (4.22) that allows to make the previous argument rigorous seems a very difficult task. So we follow a different strategy of proof, that we outline below.

For fixed $t\in[0,T]$ and $f\in H$, denote by $(Y_s^{t,f})_{s\in[t,T]}$ the mild solution to

$$\left\{\begin{array}{lcl} dY_s^{t,f}(x) &=& AY_s^{t,f}(x)\,ds+b'(s,x,X_s(x),u_s)Y_s^{t,f}(x)\,ds+\sigma_j'(s,x,X_s(x),u_s)Y_s^{t,f}(x)\,dW_s^j,\\ Y_t^{t,f}(x) &=& f(x)\end{array}\right.$$

This equation has to be compared with (4.1).

Then taking $g\in L^2$, applying the Ito formula to $\langle P_sY_s^{t,f},Y_s^{t,g}\rangle_{L^2}$ over the interval $[t,T]$, integrating from t to T and taking conditional expectation given $\mathcal{F}_t$ we formally obtain

$$\begin{aligned}\langle P_tf,g\rangle_{L^2} &= \mathbb{E}^{\mathcal{F}_t}\int_t^T\langle H_sY_s^{t,f},Y_s^{t,g}\rangle_{L^2}\,ds+\mathbb{E}^{\mathcal{F}_t}\langle hY_T^{t,f},Y_T^{t,g}\rangle_{L^2}\\ &= \mathbb{E}^{\mathcal{F}_t}\int_t^T\int_D\bar{H}(s,x)Y_s^{t,f}(x)Y_s^{t,g}(x)\,m(dx)\,ds+\mathbb{E}^{\mathcal{F}_t}\int_D\bar{h}(x)Y_T^{t,f}(x)Y_T^{t,g}(x)\,m(dx).\end{aligned}\tag{4.23}$$

The interesting fact is that this formula can be used to *define* P_t: more precisely, in Proposition 5.2 below, we will prove that if $f \in L^4$ then $(Y_s^{t,f})_{s\in[t,T]}$ is a progressive process with values in L^4, satisfying

$$\sup_{s\in[t,T]} (\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,f}\|_4^4)^{1/4} \le C\|f\|_4.$$

As a consequence we will show that the right-hand side of (4.23) defines a continuous bilinear form on L^4 (or equivalently a linear bounded operator from L^4 to $L^{4/3} = (L^4)^*$) and we will set, for $f, g \in L^4$,

$$\langle P_t f, g\rangle = \mathbb{E}^{\mathcal{F}_t}\int_t^T \int_D \bar{H}(s,x) Y_s^{t,f}(x) Y_s^{t,g}(x)\, m(dx)\, ds + \mathbb{E}^{\mathcal{F}_t}\int_D \bar{h}(x) Y_T^{t,f}(x) Y_T^{t,g}(x)\, m(dx).$$

Note that no reference to the BSDE (4.22) is needed to give this definition. Finally, it turns out that (4.21) can be proved to hold with this definition of P_t.

5 End of the proof of the stochastic maximum principle

As explained above, we are going to introduce the second adjoint process, an appropriate operator-valued process (P_t) that will allow us to conclude the proof of Theorem 3.3.

Throughout this section we assume that Hypotheses 2.1, 3.1 and 3.2 are satisfied.

The symbols $(e^{tA})_{t\ge 0}$ and A will also denote the restriction of the semigroup to the space L^4 and its infinitesimal generator in L^4. We need to recall some standard facts and constructions on analytic semigroups: see for instance [12] or [9]. Without loss of generality we can assume that A is boundedly invertible (if not, we replace A by $A - cI$ for sufficiently large constant $c > 0$ and we modify the drift coefficient accordingly). The domain of A is endowed with the norm $\|f\|_{D(A)} := \|Af\|_4$. By the analyticity assumption, one can define the fractional powers $(-A)^\eta$ of $-A$ in a standard way, for every $\eta \in (0,1)$. Each fractional power is a linear, in general unbounded, operator in L^4, with domain denoted $D(-A)^\eta$. Endowed with the norm $\|f\|_{D(-A)^\eta} := \|(-A)^\eta f\|_4$, each space $D(-A)^\eta$ is a Banach space and we have the continuous embeddings

$$D(A) \subset D(-A)^\eta \subset D(-A)^\rho \subset L^4, \qquad 0 < \rho < \eta < 1.$$

By analyticity, $e^{tA}(L^4) \subset D(A)$ for every $t > 0$, and for every $0 < \eta < 1$ there exist constants $C_1, C_\eta > 0$ such that for every $f \in L^4$ and $t \in (0,T]$,

$$\|e^{tA}f\|_{D(A)} = \|Ae^{tA}f\|_4 \le \frac{C_1}{t}\|f\|_4, \qquad \|e^{tA}f\|_{D(-A)^\eta} = \|(-A)^\eta e^{tA} f\|_4 \le \frac{C_\eta}{t^\eta}\|f\|_4.$$

Finally, as a consequence of the compact embedding $D(A) \subset L^4$, every embedding $D(-A)^\eta \subset L^4$ is also compact, $0 < \eta < 1$.

Remark 5.1 In most of what follows, we will only use the estimate $\|(-A)^\eta e^{tA}f\|_4 \le \frac{C_\eta}{t^\eta}\|f\|_4$ and the compact embedding $D(-A)^\eta \subset L^4$ for one, sufficiently small value of $\eta > 0$. This might eventually lead to a weakening of Hypothesis 3.2, but we not discuss those extensions in this paper.

For fixed $t \in [0,T]$ and $f \in L^4$, we consider the stochastic PDE

$$\left\{\begin{array}{rcl} dY_s^{t,f}(x) &=& AY_s^{t,f}(x)\, ds + b'(s,x,X_s(x),u_s) Y_s^{t,f}(x)\, ds + \sigma_j'(s,x,X_s(x),u_s) Y_s^{t,f}(x)\, dW_s^j, \\ Y_t^{t,f}(x) &=& f(x). \end{array}\right. \tag{5.1}$$

As a special case of Proposition 2.2 (with $p=4$), for every $t\in[0,T]$ there exists a unique mild solution, i.e. an adapted process $(Y_s^{t,f})_{s\in[t,T]}$ with continuous trajectories in L^4, satisfying $\mathbb{P}$-a.s.

$$Y_s^{t,f}=e^{(s-t)A}f+\int_t^s e^{(s-r)A}b'(r,\cdot,X_r(\cdot),u_r)Y_r^{t,f}(\cdot)\,dr,+\int_t^s e^{(s-r)A}\sigma_j'(r,\cdot,X_r(\cdot),u_r)Y_r^{t,f}(\cdot)\,dW_r^j,$$

for every $s\in[t,T]$. In addition we have $\sup_{0\le t\le s\le T}\mathbb{E}\|Y_s^{t,f}\|_4^4<\infty$.

Proposition 5.2 *There exists a constant C such that for $f\in L^4$, $0\le t\le s\le T$*

$$(\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,f}\|_4^4)^{1/4}\le C\|f\|_4,\qquad \mathbb{P}-a.s. \tag{5.2}$$

and for $0\le t\le t+h\le s\le T$

$$(\mathbb{E}\|Y_s^{t+h,f}-Y_s^{t,f}\|_4^4)^{1/4}\le C[\sup_{t\in[0,T]}\|(e^{tA}-e^{(t+h)A})f\|_4+h^{1/2}\|f\|_4]. \tag{5.3}$$

Moreover for every $\eta\in(0,1/4)$ there exists a constant C_η such that for $f\in D(-A)^\eta\subset L^4$, $0\le t<s\le T$

$$(\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,(-A)^\eta f}\|_4^4)^{1/4}\le C_\eta(s-t)^{-\eta}\|f\|_4,\qquad \mathbb{P}-a.s. \tag{5.4}$$

We notice that the above relation indicates that equation (5.1) regularizes the initial data (roughly speaking sends data in $D(-A)^{-\eta}$ to L^4.

Proof. For brevity we write the proof in the case $b\equiv 0$ and denote by $C_j(r)$ the multiplication operator in L^4 by the (bounded) function $\sigma_j'(r,\cdot,X_r(\cdot),u_r)$. So the equation for $Y^{t,x}$ is

$$Y_s^{t,f}=e^{(s-t)A}f+\int_t^s e^{(s-r)A}C_j(r)Y_r^{t,f}\,dW_r^j,\qquad s\in[t,T].$$

Using the conditional inequality (A.3) for $p=4$ we obtain

$$\mathbb{E}^{\mathcal{F}_t}\left\|\int_t^s e^{(s-r)A}C_j(r)Y_r^{t,f}\,dW_r^j\right\|_4^4\le C\int_t^s\mathbb{E}^{\mathcal{F}_t}\|e^{(s-r)A}C_j(r)Y_r^{t,f}\|_4^4\,dr\le C\int_t^s\mathbb{E}^{\mathcal{F}_t}\|Y_r^{t,f}\|_4^4\,dr,$$

and since $\|e^{(s-t)A}f\|_4^4\le C\|f\|_4^4$ it follows that for every $s\in[t,T]$ we have, $\mathbb{P}$-a.s.

$$\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,f}\|_4^4\le C\|f\|_4^4+C\int_t^s\mathbb{E}^{\mathcal{F}_t}\|Y_r^{t,f}\|_4^4\,dr. \tag{5.5}$$

Take a dense countable set $D\subset[t,T]$. Then, $\mathbb{P}$-a.s., (5.5) holds simultaneously for every $s\in D$. Since $Y^{t,x}$ has continuous trajectories in L^4, there exists a set N with $\mathbb{P}(N)=0$ such that $s\to\|Y_s^{t,f}(\omega)\|_4^4$ is continuous on $[t,T]$ for every $\omega\notin N$. Discarding a set of $\mathbb{P}$-measure zero, and given any $s\in[t,T]$, we take a sequence $(s_n)\subset D$, $s_n\to s$ and by the conditional Fatou Lemma

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,f}\|_4^4&=\mathbb{E}^{\mathcal{F}_t}\liminf_{n\to\infty}\|Y_{s_n}^{t,f}\|_4^4\le\liminf_{n\to\infty}\mathbb{E}^{\mathcal{F}_t}\|Y_{s_n}^{t,f}\|_4^4\\&\le C\|f\|_4^4+C\liminf_{n\to\infty}\int_t^{s_n}\mathbb{E}^{\mathcal{F}_t}\|Y_r^{t,f}\|_4^4\,dr=C\|f\|_4^4+C\int_t^s\mathbb{E}^{\mathcal{F}_t}\|Y_r^{t,f}\|_4^4\,dr.\end{aligned}$$

It follows that, $\mathbb{P}$-a.s. (5.5) holds for every $s\in[t,T]$, so that (5.2) follows from a pathwise application of Gronwall's lemma.

The proof of (5.4) is very similar: by (5.5) we have

$$\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,(-A)^\eta f}\|_4^4 \le C\|(-A)^\eta e^{(s-t)A} f\|_4^4 + C\int_t^s \mathbb{E}^{\mathcal{F}_t}\|Y_r^{t,(-A)^\eta f}\|_4^4\,dr$$
$$\le C_\eta (s-t)^{-4\eta}\|f\|_4^4 + C\int_t^s \mathbb{E}^{\mathcal{F}_t}\|Y_r^{t,(-A)^\eta f}\|_4^4\,dr$$

and (5.4) follows again from a variant of Gronwall's lemma.

To prove (5.3) we first write, for $s \in [t+h, T]$,

$$Y_s^{t+h,f} - Y_s^{t,f} = (e^{(s-t-h)A} - e^{(s-t)A})f - \int_t^{t+h} e^{(s-r)A} C_j(r) Y_r^{t,f}\,dW_r^j$$
$$+ \int_{t+h}^s e^{(s-r)A} C_j(r)(Y_r^{t+h,f} - Y_r^{t,f})\,dW_r^j =: I + II + III.$$

Then we have, using (A.2) for $p = 4$,

$$\|I\|_4 \le \sup_{t\in[0,T]} \|(e^{tA} - e^{(t+h)A})f\|_4,$$

$$\mathbb{E}\|II\|_4^4 \le ch\int_t^{t+h} \mathbb{E}\|Y_r^{t,f}\|_4^4\,dr \le ch^2\|f\|_4^4,$$

$$\mathbb{E}\|III\|_4^4 \le c\int_{t+h}^s \mathbb{E}\|Y_r^{t+h,f} - Y_r^{t,f}\|_4^4\,dr.$$

Therefore

$$\mathbb{E}\|Y_s^{t+h,f} - Y_s^{t,f}\|_4^4 \le c[\sup_{t\in[0,T]} \|(e^{tA} - e^{(t+h)A})f\|_4 + h^2\|f\|_4^4] + c\int_{t+h}^s \mathbb{E}\|Y_r^{t+h,f} - Y_r^{t,f}\|_4^4\,dr$$

and (5.3) follows from Gronwall's lemma. □

Recall that we denoted by $\mathcal{L}$ the space of linear bounded operators $L^4 \to (L^4)^* = L^{4/3}$ endowed with the usual operator norm and with the Borel σ-algebra of the weak topology. The duality between $g \in L^4$ and $h \in L^{4/3}$ is denoted $\langle h, g\rangle$. We note that, by the Hölder inequality, every $H \in L^2$ can be identified with the corresponding multiplication operator, i.e. with a unique $H \in \mathcal{L}$ satisfying

$$\langle Hf, g\rangle = \int_D H(x) f(x) g(x)\, m(dx), \qquad f, g \in L^4$$

and, moreover, $\|H\|_{\mathcal{L}} \le \|H\|_2$. Similar remarks apply to $\bar{H}(t,x)$ and $\bar{h}(x)$ defined in (4.18).

The definition of the second adjoint process P, along with some of its properties, is given in the following proposition.

Proposition 5.3 *There exists a progressive process $(P_t)_{t\in[0,T]}$ with values in $\mathcal{L}$, such that for $t \in [0,T]$, $f, g \in L^4$,*

$$\langle P_t f, g\rangle = \mathbb{E}^{\mathcal{F}_t}\int_t^T \langle \bar{H}_s Y_s^{t,f}, Y_s^{t,g}\rangle\,ds + \mathbb{E}^{\mathcal{F}_t}\langle \bar{h} Y_T^{t,f}, Y_T^{t,g}\rangle, \qquad \mathbb{P}-a.s. \tag{5.6}$$

We have

$$\sup_{t\in[0,T]} \mathbb{E}\|P_t\|_{\mathcal{L}}^2 < \infty, \tag{5.7}$$

and for every $f, g \in L^4$ we have, for $\epsilon \downarrow 0$,

$$\mathbb{E}|\langle (P_{t+\epsilon} - P_t) f, g\rangle| \to 0. \tag{5.8}$$

Moreover, for every $\eta \in (0, 1/4)$ there exists a constant C_η such that for $f, g \in D(-A)^\eta \subset L^4$, $0 \le t < T$,

$$|\langle P_t(-A)^\eta f, (-A)^\eta g\rangle| \le C_\eta \|f\|_4 \|g\|_4 (T-t)^{-2\eta} \left[\left(\int_t^T \mathbb{E}^{\mathcal{F}_t} \|\bar{H}_s\|_2^2 ds \right)^{1/2} + \left(\mathbb{E}^{\mathcal{F}_t} \|\bar{h}\|_2^2 \right)^{1/2} \right], \ \mathbb{P}\text{−}a.s. \tag{5.9}$$

which immediately implies

$$\begin{aligned} &\mathbb{E} \sup \left\{ |\langle P_t(-A)^\eta f, (-A)^\eta g\rangle|^2 \ : \ f, g \in D(-A)^\eta, \|f\|_4 \le 1, \|g\|_4 \le 1 \right\} \\ &\le C_\eta (T-t)^{-4\eta} \left[\mathbb{E} \int_0^T \|\bar{H}_s\|_2^2 ds + \mathbb{E}\|\bar{h}\|_2^2 \right]. \end{aligned} \tag{5.10}$$

Remark 5.4 1. Formula (5.6) can be written more explicitly as follows:

$$\langle P_t f, g\rangle = \mathbb{E}^{\mathcal{F}_t} \int_t^T \int_D \bar{H}(s, x) Y_s^{t,f}(x) Y_s^{t,g}(x)\, m(dx)\, ds + \mathbb{E}^{\mathcal{F}_t} \int_D \bar{h}(x) Y_T^{t,f}(x) Y_T^{t,g}(x)\, m(dx).$$

Clearly, (5.6) defines uniquely (P_t) up to modification.

2. In the following, for $T \in \mathcal{L}$ we will use the notation

$$|||T||| := \sup \left\{ |\langle T(-A)^\eta f, (-A)^\eta g\rangle| \ : \ f, g \in D(-A)^\eta, \|f\|_4 \le 1, \|g\|_4 \le 1 \right\} \tag{5.11}$$

(5.10) can then be written

$$\mathbb{E}\, |||P_t|||^2 \le C_\eta (T-t)^{-4\eta} \left[\mathbb{E} \int_0^T \|\bar{H}_s\|_2^2 ds + \mathbb{E}\|\bar{h}\|_2^2 \right]. \tag{5.12}$$

3. Note that for L^4-valued, $\mathcal{F}_t$-measurable random variables F, G we have

$$\langle P_t F, G\rangle = \mathbb{E}^{\mathcal{F}_t} \int_t^T \langle \bar{H}_s Y_s^{t,F}, Y_s^{t,G}\rangle\, ds + \mathbb{E}^{\mathcal{F}_t} \langle \bar{h} Y_T^{t,F}, Y_T^{t,G}\rangle, \qquad \mathbb{P} - a.s. \tag{5.13}$$

The equality being trivial if F and G are simple random variables and easily passing to the limit.

Proof of Proposition 5.3. Fix $\eta \in (0, 1/4)$, $f, g \in D(-A)^\eta \subset L^4$, $0 \le t < T$. Using the

conditional Hölder inequality and (5.4) we have

$$
\begin{aligned}
&\left|\mathbb{E}^{\mathcal{F}_t}\int_t^T \langle \bar{H}_s Y_s^{t,(-A)^\eta f}, Y_s^{t,(-A)^\eta g}\rangle\, ds + \mathbb{E}^{\mathcal{F}_t}\langle \bar{h} Y_T^{t,(-A)^\eta f}, Y_T^{t,(-A)^\eta g}\rangle\right| \\
&\leq \mathbb{E}^{\mathcal{F}_t}\int_t^T \|\bar{H}_s\|_2 \|Y_s^{t,(-A)^\eta f}\|_4 \|Y_s^{t,(-A)^\eta g}\|_4\, ds + \mathbb{E}^{\mathcal{F}_t}[\|\bar{h}\|_2 \|Y_T^{t,(-A)^\eta f}\|_4 \|Y_T^{t,(-A)^\eta g}\|_4] \\
&\leq \int_t^T (\mathbb{E}^{\mathcal{F}_t}\|\bar{H}_s\|_2^2)^{1/2} (\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,(-A)^\eta f}\|_4^4)^{1/4} (\mathbb{E}^{\mathcal{F}_t}\|Y_s^{t,(-A)^\eta g}\|_4^4)^{1/4}\, ds \\
&\quad + (\mathbb{E}^{\mathcal{F}_t}\|\bar{h}\|_2\|^2)^{1/2} (\mathbb{E}^{\mathcal{F}_t}\|Y_T^{t,(-A)^\eta f}\|_4^4)^{1/4} (\mathbb{E}^{\mathcal{F}_t}\|Y_T^{t,(-A)^\eta g}\|_4^4)^{1/4} \\
&\leq c\|f\|_4\|g\|_4 \int_t^T (\mathbb{E}^{\mathcal{F}_t}\|\bar{H}_s\|_2^2)^{1/2} (s-t)^{-2\eta}\, ds + c\|f\|_4\|g\|_4 (\mathbb{E}^{\mathcal{F}_t}\|\bar{h}\|_2\|^2)^{1/2} (T-t)^{-2\eta} \\
&\leq c\|f\|_4\|g\|_4 \left[\left(\int_t^T \mathbb{E}^{\mathcal{F}_t}\|\bar{H}_s\|_2^2 ds\right)^{1/2} \left(\int_t^T (s-t)^{-4\eta} ds\right)^{1/2} + (\mathbb{E}^{\mathcal{F}_t}\|\bar{h}\|_2\|^2)^{1/2} (T-t)^{-2\eta}\right] \\
&\leq c\|f\|_4\|g\|_4 (T-t)^{-2\eta} \left[\left(\int_t^T \mathbb{E}^{\mathcal{F}_t}\|\bar{H}_s\|_2^2 ds\right)^{1/2} + \left(\mathbb{E}^{\mathcal{F}_t}\|\bar{h}\|_2^2\right)^{1/2}\right],
\end{aligned}
\tag{5.14}
$$

where c is a constant independent of f, g, t. Using (5.2) instead of (5.4) this inequality also holds for $\eta = 0$.

Now fix a dense set F in L^4. For $f, g \in F$ let us define $\langle P_t f, g\rangle$ by formula (5.6), by fixing an arbitrary version of the conditional expectations on the right-hand side. By (5.14) with $\eta = 0$, there exists a set N of probability zero such that for $\omega \notin F$ we have

$$
|\langle P_t(\omega) f, g\rangle| \leq c\|f\|_4\|g\|_4, \qquad f, g \in F.
$$

Thus, the mapping $(f, g) \mapsto \langle P_t(\omega) f, g\rangle$ extends from $F \times F$ to a continuous bilinear form on L^4 (or equivalently an element of $\mathcal{L}$), still denoted $P_t(\omega)$. Set $P_t(\omega) = 0$ for $\omega \in N$. Using again (5.14) with $\eta = 0$, it is easily proved that equality (5.6) holds for every $f, g \in L^4$, by approximating f, g with elements of F. Thus, an $\mathcal{L}$-valued process $(P_t)_{t\in[0,T]}$ has been constructed with the required properties. (P_t) is adapted by construction. Similar arguments also show the existence of a progressive modification of (P_t), as required.

(5.9) follows at once from (5.14). (5.14) with $\eta = 0$ gives

$$
\|P_t\| \leq c \left[\left(\int_t^T \mathbb{E}^{\mathcal{F}_t}\|\bar{H}_s\|_2^2 ds\right)^{1/2} + \left(\mathbb{E}^{\mathcal{F}_t}\|\bar{h}\|_2^2\right)^{1/2}\right],
$$

which implies (5.7).

It remains to prove (5.8). We sketch the proof in the case $\bar{h} = 0$ for short.

$$
\begin{aligned}
&\langle (P_{t+\epsilon} - P_t) f, g\rangle = (\mathbb{E}^{\mathcal{F}_{t+\epsilon}} - \mathbb{E}^{\mathcal{F}_t}) \int_t^T \langle \bar{H}_s Y_s^{t,f}, Y_s^{t,g}\rangle\, ds \\
&- \mathbb{E}^{\mathcal{F}_{t+\epsilon}} \int_t^{t+\epsilon} \langle \bar{H}_s Y_s^{t,f}, Y_s^{t,g}\rangle\, ds + \mathbb{E}^{\mathcal{F}_t} \int_{t+\epsilon}^T [\langle \bar{H}_s Y_s^{t+\epsilon,f}, Y_s^{t+\epsilon,g}\rangle - \langle \bar{H}_s Y_s^{t,f}, Y_s^{t,g}\rangle]\, ds.
\end{aligned}
$$

The first summand tends to zero in $L^1(\Omega, \mathbb{P})$ by the downwards martingale convergence theorem, the third one due to (5.3) and the second one is easy to treat by dominated convergence Theorem. □

We are now ready to finish the proof of our main result, by showing that the formula (4.21) introduced during our heuristic discussion actually holds (more precisely we will prove (5.15) below).

End of the proof of Theorem 3.3. We claim that the following holds:

$$\mathbb{E}\int_0^T \langle \bar{H}_s Y_s^\epsilon, Y_s^\epsilon\rangle\, ds + \mathbb{E}\langle \bar{h} Y_T^\epsilon, Y_T^\epsilon\rangle = \mathbb{E}\int_0^T \langle P_s \delta^\epsilon \sigma_j(s,\cdot), \delta^\epsilon \sigma_j(s,\cdot)\rangle\, ds + o(\epsilon). \tag{5.15}$$

Admitting this for a moment, if follows from Proposition 4.6 that

$$\begin{aligned} J(u^\epsilon) - J(u) &= \mathbb{E}\int_0^T \int_D [\delta^\epsilon l(t,x) + p_t(x)\delta^\epsilon b(t,x) + q_t^j(x)\delta^\epsilon \sigma_j(t,x)]\, ds\, m(dx) \\ &\quad + \mathbb{E}\int_0^T \langle P_s \delta^\epsilon \sigma_j(s,\cdot), \delta^\epsilon \sigma_j(s,\cdot)\rangle\, ds + o(\epsilon). \end{aligned}$$

The optimality of u implies that $J(u^\epsilon) - J(u) \geq 0$. Diving by ϵ and letting $\epsilon \to 0$, the required conclusion is obtained by standard arguments, see e.g. [13] or [18].

So it only remains to prove (5.15). Recalling that $Y_s^\epsilon = 0$ for $s \leq t_0$, the left-hand side of (5.15) equals

$$\mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle \bar{H}_s Y_s^\epsilon, Y_s^\epsilon\rangle\, ds + \mathbb{E}\int_{t_0+\epsilon}^T \langle \bar{H}_s Y_s^\epsilon, Y_s^\epsilon\rangle\, ds + \mathbb{E}\langle \bar{h} Y_T^\epsilon, Y_T^\epsilon\rangle.$$

It is easily checked that the first integral is $o(\epsilon)$. Using the formula

$$Y_s^\epsilon = Y_s^{t_0+\epsilon, Y_{t_0+\epsilon}^\epsilon}, \qquad s \geq t_0 + \epsilon,$$

which follows by comparing the equations (5.1) and (4.1) satisfied by $Y^{t_0+\epsilon, f}$ and Y^ϵ, we obtain

$$\begin{aligned} &\mathbb{E}\int_0^T \langle \bar{H}_s Y_s^\epsilon, Y_s^\epsilon\rangle\, ds + \mathbb{E}\langle \bar{h} Y_T^\epsilon, Y_T^\epsilon\rangle \\ &= o(\epsilon) + \mathbb{E}\int_{t_0+\epsilon}^T \langle \bar{H}_s Y_s^{t_0+\epsilon, Y_{t_0+\epsilon}^\epsilon}, Y_s^{t_0+\epsilon, Y_{t_0+\epsilon}^\epsilon}\rangle\, ds + \mathbb{E}\langle \bar{h} Y_T^{t_0+\epsilon, Y_{t_0+\epsilon}^\epsilon}, Y_T^{t_0+\epsilon, Y_{t_0+\epsilon}^\epsilon}\rangle \\ &= o(\epsilon) + \mathbb{E}\langle P_{t_0+\epsilon} Y_{t_0+\epsilon}^\epsilon, Y_{t_0+\epsilon}^\epsilon\rangle, \end{aligned} \tag{5.16}$$

where the last equality follows from an application of (5.13). Next we claim that

$$\mathbb{E}\langle (P_{t_0+\epsilon} - P_{t_0})\, Y_{t_0+\epsilon}^\epsilon, Y_{t_0+\epsilon}^\epsilon\rangle = o(\epsilon), \tag{5.17}$$

$$\mathbb{E}\langle P_{t_0} Y_{t_0+\epsilon}^\epsilon, Y_{t_0+\epsilon}^\epsilon\rangle = \mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle P_s \delta^\epsilon \sigma_j(s,\cdot), \delta^\epsilon \sigma_j(s,\cdot)\rangle\, ds + o(\epsilon). \tag{5.18}$$

The required formula (5.15) will now be a consequence of (5.17) and (5.18), which are proved in the following two subsection below. The proof of Theorem 3.3 will then be finished. □

5.1 Proof of (5.17)

It is convenient to rewrite (5.17) in the form

$$\mathbb{E}\langle (P_{t_0+\epsilon} - P_{t_0})\, \epsilon^{-1/2} Y_{t_0+\epsilon}^\epsilon, \epsilon^{-1/2} Y_{t_0+\epsilon}^\epsilon\rangle \to 0. \tag{5.19}$$

By Proposition 4.1 there exists a constant C_0 independent of ϵ such that

$$(\mathbb{E}\|\epsilon^{-1/2} Y_{t_0+\epsilon}^\epsilon\|_4^4)^{1/4} \leq C_0, \quad (\mathbb{E}\|\epsilon^{-1/2} Y_{t_0+\epsilon}^\epsilon\|_4^8)^{1/8} \leq C_0. \tag{5.20}$$

Next we fix $\eta \in (0, 1/4)$ and notice that for every $\delta > 0$ we have, by the Markov inequality,

$$\mathbb{P}(\|\epsilon^{-1/2}(-A)^{-\eta} Y_{t_0+\epsilon}^\epsilon\|_{D(-A)^\eta} > C_0\delta^{-1/4}) = \mathbb{P}(\|\epsilon^{-1/2} Y_{t_0+\epsilon}^\epsilon\|_4 > C_0\delta^{-1/4}) \leq \delta.$$

Therefore setting $K_\delta = \{f \in L^4 \ : \ f \in D(-A)^\eta, \|f\|_{D(-A)^\eta} \le C_0\delta^{-1/4}\}$ and denoting $\Omega_{\delta,\epsilon}$ the event $\{\epsilon^{-1/2}(-A)^{-\eta}Y^\epsilon_{t_0+\epsilon} \in K_\delta\}$ we obtain

$$\mathbb{P}(\Omega^c_{\delta,\epsilon}) = \mathbb{P}(\epsilon^{-1/2}(-A)^{-\eta}Y^\epsilon_{t_0+\epsilon} \notin K_\delta) \le \delta.$$

We note that, since $D(-A)^\eta$ is compactly embedded in L^4, the set K_δ is a compact subset of L^4. Moreover, for $f \in K_\delta$ we have

$$\|f\|_4 \le c\|f\|_{D(-A)^\eta} \le cC_0\delta^{-1/4}, \tag{5.21}$$

i.e. K_δ is contained in a ball of L^4 centered at 0 with radius proportional to $\delta^{-1/4}$.

We have

$$\begin{aligned}
&\mathbb{E}\langle(P_{t_0+\epsilon} - P_{t_0})\,\epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}, \epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}\rangle \\
&= \mathbb{E}[\langle(P_{t_0+\epsilon} - P_{t_0})\,\epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}, \epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}\rangle 1_{\Omega^c_{\delta,\epsilon}}] + \mathbb{E}[\langle(P_{t_0+\epsilon} - P_{t_0})\,\epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}, \epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}\rangle 1_{\Omega_{\delta,\epsilon}}] \\
&=: A^\epsilon_1 + A^\epsilon_2.
\end{aligned}$$

By the Hölder inequality

$$|A^\epsilon_1| \le (\mathbb{E}\|P_{t_0+\epsilon} - P_{t_0}\|^2_{\mathcal{L}})^{1/2}(\mathbb{E}\|\epsilon^{-1/2}Y^\epsilon_{t_0+\epsilon}\|^8_4)^{1/4}\mathbb{P}(\Omega^c_{\delta,\epsilon})^{1/4},$$

and from (5.7), (5.20) we conclude that $|A^\epsilon_1| \le c\mathbb{P}(\Omega^c_{\delta,\epsilon})^{1/4} \le c\delta^{1/4}$ for some constant c independent of δ and ϵ.

On the other hand, recalling the definition of $\Omega_{\delta,\epsilon}$,

$$|A^\epsilon_2| \le \mathbb{E}\sup_{f\in K_\delta} |\langle(P_{t_0+\epsilon} - P_{t_0})\,(-A)^\eta f, (-A)^\eta f\rangle 1_{\Omega_{\delta,\epsilon}}|.$$

Since K_δ is compact in L^4, it can be covered by a finite number N_δ of open balls with radius δ and centers denoted f^δ_i, $i = 1, \ldots, N_\delta$. Since $D(-A)^\eta$ is dense in L^4, we can assume that $f^\delta_i \in D(-A)^\eta$. Given $f \in K_\delta$, let i be such that $\|f - f^\delta_i\|_4 < \delta$; then writing

$$\begin{aligned}
&\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f, (-A)^\eta f\rangle = \langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f^\delta_i, (-A)^\eta f^\delta_i\rangle \\
&-\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta (f - f^\delta_i), (-A)^\eta (f - f^\delta_i)\rangle + 2\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f, (-A)^\eta (f - f^\delta_i)\rangle
\end{aligned}$$

and recalling the notation introduced in (5.11) we obtain

$$\begin{aligned}
&|\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f, (-A)^\eta f\rangle| \le |\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f^\delta_i, (-A)^\eta f^\delta_i\rangle| \\
&+|||P_{t_0+\epsilon} - P_{t_0}|||\,\delta^2 + 2|||P_{t_0+\epsilon} - P_{t_0}|||\,\|f\|_4\,\delta.
\end{aligned}$$

Recalling (5.21) we conclude that

$$\begin{aligned}
&\sup_{f\in K_\delta} |\langle(P_{t_0+\epsilon} - P_{t_0})\,(-A)^\eta f, (-A)^\eta f\rangle \le \sum_{i=1}^{N_\delta} |\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f^\delta_i, (-A)^\eta f^\delta_i\rangle| \\
&+2\sup_{t\in[t_0,t_0+\epsilon]} |||P_t|||\,\delta^2 + c\sup_{t\in[t_0,t_0+\epsilon]} |||P_t|||\,\delta^{3/4},
\end{aligned}$$

for some constant c. Taking expectation, it follows from (5.12) that

$$|A^\epsilon_2| \le \sum_{i=1}^{N_\delta} \mathbb{E}|\langle(P_{t_0+\epsilon} - P_{t_0})(-A)^\eta f^\delta_i, (-A)^\eta f^\delta_i\rangle| + c(T - t_0 - \epsilon)^{-2\eta}[\delta^2 + \delta^{3/4}],$$

for some constant c independent of ϵ and δ. By (5.8) we conclude that

$$\limsup_{\epsilon\downarrow 0} |A^\epsilon_2| \le c(T - t_0)^{-2\eta}[\delta^2 + \delta^{3/4}].$$

Letting $\delta \to 0$ we obtain $|A^\epsilon_1| + |A^\epsilon_2| \to 0$ and the proof of (5.17) is finished.

5.2 Proof of (5.18)

In order to make appropriate computations on $\mathbb{E}\langle P_{t_0}\, Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle$ we perform an approximation of both P_{t_0} and $Y^\epsilon_{t_0+\epsilon}$.

To approximate P_{t_0} we use the basis $(e_i)_{i\geq 1}$ of Hypothesis 3.1. We introduce the projection operators $\Pi_N f = \sum_{i=1}^N \langle f, e_i\rangle_2 e_i$, $f \in L^2$, where $\langle \cdot, \cdot\rangle_2$ denotes the scalar product of L^2. Each Π_N is an orthogonal projection in L^2. Since we assume that $(e_i)_{i\geq 1}$ is a Schauder basis of L^4, the restriction of Π_N to L^4 is a bounded linear operator in L^4, satisfying $\|\Pi_N f - f\|_4 \to 0$ for every $f \in L^4$ and $\sup_N \|\Pi_N\|_{L(L^4,L^4)} < \infty$. Then we define

$$P_t^N(\omega) f := \sum_{i,j=1}^N \langle P_t(\omega) e_i, e_j\rangle \langle e_i, f\rangle_2 e_j, \qquad f \in L^4.$$

Then $P_t^N(\omega)$ is a linear bounded operator on L^4, which extends to a linear bounded operator on L^2, with values in the finite-dimensional subspace spanned by $e_1, \ldots, e_N$. Moreover

$$\langle P_t^N(\omega) f, g\rangle_2 = \sum_{i,j=1}^N \langle P_t(\omega) e_i, e_j\rangle \langle e_i, f\rangle_2 \langle e_j, g\rangle_2 = \langle P_t(\omega)\Pi_N f, \Pi_N g\rangle, \qquad f, g \in L^4. \tag{5.22}$$

In the following we will consider P^N as a stochastic process with values in $\mathcal{L}_2(L^2)$, the space of Hilbert-Schmidt operators on L^2.

In order to approximate $Y^\epsilon_{t_0+\epsilon}$ we introduce

$$J_n = (nI - A)^{-1}, \qquad A_n = A J_n, \qquad Y_t^{\epsilon,n}(\omega) = J_n Y_t^\epsilon(\omega).$$

Note that A_n are the Yosida approximations of the operator A.

We are going to approximate $\mathbb{E}\langle P_{t_0}\, Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle$ by $\mathbb{E}\langle P^N_{t_0}\, Y^{\epsilon,n}_{t_0+\epsilon}, Y^{\epsilon,n}_{t_0+\epsilon}\rangle_2$.

$Y^{\epsilon,n}$ is a process with values in L^2 which admits an Ito differential that we are going to compute. Recall equation (4.2) satisfied by Y^ϵ, that we now re-write in the following way: for $s \geq t_0$,

$$Y_s^\epsilon = \int_{t_0}^s e^{(s-r)A}[B(r) Y_r^\epsilon + \delta^\epsilon b(r)]\, dr + \int_{t_0}^s e^{(s-r)A}[C_j(r) Y_r^\epsilon + \delta^\epsilon \sigma_j(r)]\, dW_r^j, \qquad \mathbb{P} - a.s.$$

where $B(r), C_j(r)$ denote the multiplication operators by the functions $b'(r, \cdot, X_r(\cdot), u_r)$ and $\sigma_j'(r, \cdot, X_r(\cdot), u_r)$ respectively. Applying J_n to both sides it is not hard to conclude that $Y_t^{\epsilon,n}$ has the Ito differential

$$dY_s^{\epsilon,n} = A_n\, Y_s^{\epsilon,n}\, ds + [J_n B(s) Y_s^\epsilon + J_n \delta^\epsilon b(s)]\, ds + [J_n C_j(s) Y_s^\epsilon + J_n \delta^\epsilon \sigma_j(s)]\, dW_s^j.$$

In the following for $y, z \in L^2$, we denote by $y \otimes z$ the rank-one operator $f \mapsto \langle f, z\rangle_2\, y$ on L^2. Using this notation we will consider the $\mathcal{L}_2(L^2)$-valued process $Y_s^{\epsilon,n} \otimes Y_s^{\epsilon,n}$, $s \in [t_0, T]$ (recall that if K is a separable Hilbert space, $\mathcal{L}_2(K)$ is the Hilbert space of all bounded linear operators in X for which $||X||^2_{\mathcal{L}_2(K)} = tr(X^*X)$ is finite naturally endowed with the product $\langle X_1, X_2\rangle_{\mathcal{L}_2(K)} = tr(X_1^* X_2)$).

By the Ito formula for Hilbert-space valued Ito processes we have

$$\begin{aligned}
d(Y_s^{\epsilon,n} \otimes Y_s^{\epsilon,n}) &= A_n\, (Y_s^{\epsilon,n} \otimes Y_s^{\epsilon,n})\, ds + (Y_s^{\epsilon,n} \otimes Y_s^{\epsilon,n})\, A_n^*\, ds \\
&\quad + Y_s^{\epsilon,n} \otimes [J_n B(s) Y_s^\epsilon + J_n \delta^\epsilon b(s)]\, ds + [J_n B(s) Y_s^\epsilon + J_n \delta^\epsilon b(s)] \otimes Y_s^{\epsilon,n}\, ds \\
&\quad + Y_s^{\epsilon,n} \otimes [J_n C_j(s) Y_s^\epsilon + J_n \delta^\epsilon \sigma_j(s)]\, dW_s^j + [J_n C_j(s) Y_s^\epsilon + J_n \delta^\epsilon \sigma_j(s)] \otimes Y_s^{\epsilon,n}\, dW_s^j \\
&\quad + [J_n C_j(s) Y_s^\epsilon + J_n \delta^\epsilon \sigma_j(s)] \otimes [J_n C_j(s) Y_s^\epsilon + J_n \delta^\epsilon \sigma_j(s)]\, ds
\end{aligned}$$

and it follows that

$$
\begin{aligned}
&Y^{\epsilon,n}_s \otimes Y^{\epsilon,n}_s\\
&= \int_{t_0}^{s} e^{(s-r)A_n}\{Y^{\epsilon,n}_r \otimes [J_n B(r) Y^\epsilon_r + J_n \delta^\epsilon b(r)] + [J_n B(r) Y^\epsilon_r + J_n \delta^\epsilon b(r)] \otimes Y^{\epsilon,n}_r\} e^{(s-r)A_n^*}\, dr\\
&+ \int_{t_0}^{s} e^{(s-r)A_n}\{Y^{\epsilon,n}_r \otimes [J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)] + [J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)] \otimes Y^{\epsilon,n}_r\} e^{(s-r)A_n^*}\, dW^j_r\\
&+ \int_{t_0}^{s} e^{(s-r)A_n}\{[J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)] \otimes [J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)]\} e^{(s-r)A_n^*}\, dr
\end{aligned}
$$

The reason for introducing the process $Y^{\epsilon,n} \otimes Y^{\epsilon,n}$ is that we can now make the following computation: denoting by tr the trace of operators in L^2 we have

$$
\mathbb{E}\langle P^N_{t_0}\, Y^{\epsilon,n}_{t_0+\epsilon}, Y^{\epsilon,n}_{t_0+\epsilon}\rangle_2 = \mathbb{E}\, tr[P^N_{t_0}\, (Y^{\epsilon,n}_{t_0+\epsilon} \otimes Y^{\epsilon,n}_{t_0+\epsilon})]
$$

and we can replace $(Y^{\epsilon,n}_{t_0+\epsilon} \otimes Y^{\epsilon,n}_{t_0+\epsilon})$ by the previous formula. Taking conditional expectation with respect to $\mathcal{F}_{t_0}$ the stochastic integral disappears and we obtain

$$
\begin{aligned}
&\mathbb{E}\langle P^N_{t_0}\, Y^{\epsilon,n}_{t_0+\epsilon}, Y^{\epsilon,n}_{t_0+\epsilon}\rangle_2\\
&= \int_{t_0}^{t_0+\epsilon} \mathbb{E}\, tr\Big[P^N_{t_0}\, e^{(t_0+\epsilon-r)A_n}\{Y^{\epsilon,n}_r \otimes [J_n B(r) Y^\epsilon_r + J_n \delta^\epsilon b(r)]\\
&+ [J_n B(r) Y^\epsilon_r + J_n \delta^\epsilon b(r)] \otimes Y^{\epsilon,n}_r\} e^{(t_0+\epsilon-r)A_n^*}\Big]\, dr\\
&+ \int_{t_0}^{t_0+\epsilon} \mathbb{E}\, tr\Big[P^N_{t_0}\, e^{(t_0+\epsilon-r)A_n}\{[J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)] \otimes [J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)]\} e^{(t_0+\epsilon-r)A_n^*}\Big]\, dr\\
&= 2\mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P^N_{t_0}\, e^{(t_0+\epsilon-r)A_n}[J_n B(r) Y^\epsilon_r + J_n \delta^\epsilon b(r)], e^{(t_0+\epsilon-r)A_n} Y^{\epsilon,n}_r\rangle_2\, dr\\
&+ \mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P^N_{t_0}\, e^{(t_0+\epsilon-r)A_n}[J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)], e^{(t_0+\epsilon-r)A_n}[J_n C_j(r) Y^\epsilon_r + J_n \delta^\epsilon \sigma_j(r)]\rangle_2\, dr.
\end{aligned}
$$

Next we let $n \to \infty$ and we use the fact that $\|e^{tA_n} f - e^{tA} f\|_2 \to 0$ and $\|J_n f - f\|_2 \to 0$ for $f \in L^2$. It follows that

$$
\begin{aligned}
&\mathbb{E}\langle P^N_{t_0}\, Y^{\epsilon}_{t_0+\epsilon}, Y^{\epsilon}_{t_0+\epsilon}\rangle_2 = 2\mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P^N_{t_0}\, e^{(t_0+\epsilon-r)A}[B(r) Y^\epsilon_r + \delta^\epsilon b(r)], e^{(t_0+\epsilon-r)A} Y^{\epsilon}_r\rangle_2\, dr\\
&+ \mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P^N_{t_0}\, e^{(t_0+\epsilon-r)A}[C_j(r) Y^\epsilon_r + \delta^\epsilon \sigma_j(r)], e^{(t_0+\epsilon-r)A}[C_j(r) Y^\epsilon_r + \delta^\epsilon \sigma_j(r)]\rangle_2\, dr.
\end{aligned}
$$

Recalling (5.22), this formula can be written

$$
\begin{aligned}
&\mathbb{E}\langle P_{t_0} \Pi^N Y^{\epsilon}_{t_0+\epsilon}, \Pi^N Y^{\epsilon}_{t_0+\epsilon}\rangle = 2\mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P_{t_0} \Pi^N e^{(t_0+\epsilon-r)A}[B(r) Y^\epsilon_r + \delta^\epsilon b(r)], \Pi^N e^{(t_0+\epsilon-r)A} Y^{\epsilon}_r\rangle\, dr\\
&+ \mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P_{t_0} \Pi^N e^{(t_0+\epsilon-r)A}[C_j(r) Y^\epsilon_r + \delta^\epsilon \sigma_j(r)], \Pi^N e^{(t_0+\epsilon-r)A}[C_j(r) Y^\epsilon_r + \delta^\epsilon \sigma_j(r)]\rangle\, dr.
\end{aligned}
$$

We let $N \to \infty$ and we finally obtain

$$
\begin{aligned}
&\mathbb{E}\langle P_{t_0}\, Y^{\epsilon}_{t_0+\epsilon}, Y^{\epsilon}_{t_0+\epsilon}\rangle = 2\mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P_{t_0}\, e^{(t_0+\epsilon-r)A}[B(r) Y^\epsilon_r + \delta^\epsilon b(r)], e^{(t_0+\epsilon-r)A} Y^{\epsilon}_r\rangle\, dr\\
&+ \mathbb{E} \int_{t_0}^{t_0+\epsilon} \langle P_{t_0}\, e^{(t_0+\epsilon-r)A}[C_j(r) Y^\epsilon_r + \delta^\epsilon \sigma_j(r)], e^{(t_0+\epsilon-r)A}[C_j(r) Y^\epsilon_r + \delta^\epsilon \sigma_j(r)]\rangle\, dr.
\end{aligned}
$$

Using the estimate in Proposition 4.1 it follows that

$$\mathbb{E}\langle P_{t_0}\, Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle = \mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle P_{t_0}\, e^{(t_0+\epsilon-r)A}\delta^\epsilon\sigma_j(r), e^{(t_0+\epsilon-r)A}\delta^\epsilon\sigma_j(r)\rangle\, dr + o(\epsilon),$$

and since $\|e^{tA}f - f\|_4 \to 0$ as $t \to 0$ for every $f \in L^4$ we also conclude that

$$\mathbb{E}\langle P_{t_0}\, Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle = \mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle P_{t_0}\, \delta^\epsilon\sigma_j(r), \delta^\epsilon\sigma_j(r)\rangle\, dr + o(\epsilon).$$

Therefore, in order to finish the proof of (5.18), it remains to show that

$$\mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle (P_r - P_{t_0})\, \delta^\epsilon\sigma_j(r), \delta^\epsilon\sigma_j(r)\rangle\, dr = o(\epsilon). \tag{5.23}$$

We fix $\eta \in (0, 1/4)$. Since we have $\|\delta^\epsilon\sigma_j(s)\|_4 \le C_0$, for some constant C_0, it follows that

$$(-A)^{-\eta}\delta^\epsilon\sigma_j(s) \in K := \{f \in L^4 \ : \ f \in D(-A)^\eta, \|f\|_{D(-A)^\eta} \le C_0\}.$$

Since $D(-A)^\eta$ is compactly embedded in L^4, the set K is a compact, hence bounded, subset of L^4. We have

$$|\mathbb{E}\langle (P_r - P_{t_0})\, \delta^\epsilon\sigma_j(r), \delta^\epsilon\sigma_j(r)\rangle| \le \mathbb{E}\sup_{f\in K} |\langle (P_r - P_{t_0})\, (-A)^\eta f, (-A)^\eta f\rangle|.$$

Since K is compact in L^4, for every $\delta > 0$ it can be covered by a finite number N_δ of open balls with radius δ and centers denoted f_i^δ, $i = 1, \dots, N_\delta$. Since $D(-A)^\eta$ is dense in L^4, we can assume that $f_i^\delta \in D(-A)^\eta$. Given $f \in K$, let i be such that $\|f - f_i^\delta\|_4 < \delta$; then writing

$$\begin{aligned}&\langle (P_r - P_{t_0})(-A)^\eta f, (-A)^\eta f\rangle = \langle (P_r - P_{t_0})(-A)^\eta f_i^\delta, (-A)^\eta f_i^\delta\rangle \\ &-\langle (P_r - P_{t_0})(-A)^\eta (f - f_i^\delta), (-A)^\eta (f - f_i^\delta)\rangle + 2\langle (P_r - P_{t_0})(-A)^\eta f, (-A)^\eta (f - f_i^\delta)\rangle\end{aligned}$$

and recalling the notation introduced in (5.11) we obtain

$$\begin{aligned}&|\langle (P_r - P_{t_0})(-A)^\eta f, (-A)^\eta f\rangle| \le |\langle (P_r - P_{t_0})(-A)^\eta f_i^\delta, (-A)^\eta f_i^\delta\rangle| \\ &+|||P_r - P_{t_0}|||\, \delta^2 + 2|||P_r - P_{t_0}|||\, \|f\|_4\, \delta.\end{aligned}$$

Since K is bounded in L^4, we conclude that

$$\begin{aligned}&\sup_{f\in K} |\langle (P_r - P_{t_0})\, (-A)^\eta f, (-A)^\eta f\rangle \le \sum_{i=1}^{N_\delta} |\langle (P_r - P_{t_0})(-A)^\eta f_i^\delta, (-A)^\eta f_i^\delta\rangle| \\ &+2 \sup_{t\in[t_0,r]} |||P_t|||\, \delta^2 + c \sup_{t\in[t_0,r]} |||P_t|||\, \delta,\end{aligned}$$

for some constant c. Taking expectation, it follows from (5.12) that

$$|\mathbb{E}\langle (P_r - P_{t_0})\, \delta^\epsilon\sigma_j(r), \delta^\epsilon\sigma_j(r)\rangle| \le \sum_{i=1}^{N_\delta} \mathbb{E}|\langle (P_r - P_{t_0})(-A)^\eta f_i^\delta, (-A)^\eta f_i^\delta\rangle| + c(T - r)^{-2\eta}[\delta^2 + \delta],$$

for some constant c independent of ϵ and δ. By (5.8) we conclude that

$$\limsup_{r\downarrow t_0} |\mathbb{E}\langle (P_r - P_{t_0})\, \delta^\epsilon\sigma_j(r), \delta^\epsilon\sigma_j(r)\rangle| \le c(T - t_0)^{-2\eta}[\delta^2 + \delta].$$

Letting $\delta \to 0$ we conclude that the left-hand side is zero, and (5.23) follows immediately.

A Stochastic integrals in L^p spaces

In this appendix we sketch the construction and some basic properties of stochastic integrals with respect to a finite dimensional Wiener process, taking values in an L^p-space. The few facts collected below are enough for the present paper.

Let $(W_t^1, \dots, W_t^d)_{t\geq 0}$ be a standard, d-dimensional Wiener process defined in some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $(\mathcal{F}_t)_{t\geq 0}$ the corresponding natural filtration, augmented in the usual way, and we denote by $\mathcal{P}$ the progressive σ-algebra on $\Omega \times [0,T]$, where $T > 0$ is a given number. Let $L^p := L^p(D, \mathcal{D}, m)$ be the usual space, where m is a positive, σ-finite measure and $p \in [2, \infty)$. The integrand processes will be functions $H : \Omega \times [0,T] \times D \to \mathbb{R}^d$, which are assumed to be $\mathcal{P} \otimes \mathcal{D}$-measurable. When H is of special type, i.e. it has components of the form

$$H^j(\omega, t, x) = \sum_{i=1}^N h_i^j(\omega, t) f_i^j(x)$$

for $j = 1, \dots, d$, h_i^j bounded $\mathcal{P}$-measurable, f_i^j bounded $\mathcal{D}$-measurable, then the stochastic integral $I_t(x)$ is defined for fixed $x \in D$ by the formula $I_t(x) = \int_0^t H_s^j(x)\, dW_s^j = f_i^j(x) \int_0^t h_i^j(s)\, dW_s^j$. Using the Burkholder-Davis-Gundy inequalities for real-valued stochastic integrals, we have for some constant c_p (depending only on p):

$$\mathbb{E}|I_t(x)|^p \leq c_p \mathbb{E} \left(\int_0^t |H_s(x)|^2 ds \right)^{p/2}$$

where $|H_s(x)|^2 = \sum_{j=1}^d |H_s^j(x)|^2$. Since $p \geq 2$ we have, by en elementary inequality,

$$\mathbb{E}|I_t(x)|^p \leq c_p \left(\int_0^t (\mathbb{E}|H_s(x)|^p)^{2/p} ds \right)^{p/2} = c_p \left(\int_0^t \|H_s(x)\|^2_{L^p(\Omega;\mathbb{R}^d)} ds \right)^{p/2}.$$

Integrating with respect to m we obtain, again by elementary arguments,

$$\mathbb{E}\|I_t\|^p_{L^p(D)} \leq c_p \int_D \left(\int_0^t (\mathbb{E}|H_s(x)|^p)^{2/p} ds \right)^{p/2} m(dx) \leq c_p \left(\int_0^t \left(\int_D \mathbb{E}|H_s(x)|^p m(dx) \right)^{2/p} ds \right)^{p/2}$$

which can be written

$$\mathbb{E}\|I_t\|^p_{L^p(D)} \leq c_p \left(\int_0^t (\mathbb{E}\|H_s\|^p_{L^p(D;\mathbb{R}^d)})^{2/p} ds \right)^{p/2} \tag{A.1}$$

or equivalently

$$\|I_t\|_{L^p(\Omega \times D)} \leq c_p^{1/p} \left(\int_0^t \|H_s\|^2_{L^p(\Omega \times D;\mathbb{R}^d)} ds \right)^{1/2}.$$

Finally, by standard arguments, the stochastic integral can be extended to the class of $\mathcal{P} \otimes \mathcal{D}$-measurable integrands H for which the right-hand side of (A.1) is finite, and the inequality (A.1) remains true.

We finally note that from (A.1) and the Hölder inequality it follows that

$$\mathbb{E}\|I_t\|^p_{L^p(D)} \leq c_p \int_0^t \mathbb{E}\|H_s\|^p_{L^p(D;\mathbb{R}^d)} ds\; t^{(p-2)/2}. \tag{A.2}$$

Now suppose that there exist regular conditional probabilities $\mathbb{P}(\cdot|\mathcal{F}_t)$ given any $\mathcal{F}_t$ (this holds for instance if the Wiener process is canonically realized on the space of $\mathbb{R}^d$-valued continuous

functions). Then a slight modification of the previous passages shows the validity of the following conditional variant of (A.2): for $0 \leq r \leq t$,

$$\mathbb{E}^{\mathcal{F}_r} \| \int_r^t H_s^j \, dW_s^j \|^p_{L^p(D)} \leq c_p \int_r^t \mathbb{E}^{\mathcal{F}_r} \|H_s\|^p_{L^p(D;\mathbb{R}^d)} ds \; (t-r)^{(p-2)/2}. \tag{A.3}$$

This is used in the proof of Proposition 5.2.

References

[1] A. Bensoussan. Stochastic maximum principle for distributed parameter systems. J. Franklin Inst. 315 (1983), no. 5-6, 387–406.

[2] G. Da Prato, J. Zabczyk. Stochastic equations in infinite dimensions. Encyclopedia of Mathematics and its Applications, 44. Cambridge University Press, Cambridge, 1992.

[3] K. Du, Q. Meng. Stochastic maximum principle for infinite dimensional control systems. Preprint arXiv:1208.0529.

[4] K. Du, Q. Meng. A General Maximum Principle for Optimal Control of Stochastic Evolution Equations. Preprint arXiv:1206.3649.

[5] M. Fuhrman, Y. Hu, G. Tessitore. Stochastic maximum principle for optimal control of SPDEs. C. R. Math. Acad. Sci. Paris 350 (2012), no. 13-14, 683–688.

[6] G. Guatteri. Stochastic maximum principle for SPDEs with noise and control on the boundary. Systems Control Lett. 60 (2011), no. 3, 198–204.

[7] Y. Hu, S. Peng. Adapted solution of a backward semilinear stochastic evolution equation. Stochastic Anal. Appl. 9 (1991), no. 4, 445–459.

[8] Y. Hu, S. Peng. Maximum principle for semilinear stochastic evolution control systems. Stochastics Stochastics Rep. 33 (1990), no. 3-4, 159–180.

[9] A. Lunardi. Analytic semigroups and optimal regularity in parabolic problems. Progress in Nonlinear Differential Equations and their Applications, 16. Birkhauser, Basel, 1995.

[10] Q. Lü, X. Zhang. General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions. Preprint arXiv:1204.3275.

[11] B. Øksendal, A. Sulem, T. Zhang. Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations. Adv. in Appl. Probab. 43 (2011), no. 2, 572–596.

[12] A. Pazy. Semigroups of linear operators and applications to partial differential equations. Applied Mathematical Sciences, 44. Springer, New York, 1983.

[13] S. Peng. A general stochastic maximum principle for optimal control problems. SIAM J. Control Optim. 28 (1990), no. 4, 966–979.

[14] S. Tang, X. Li. Maximum principle for optimal control of distributed parameter stochastic systems with random jumps. Differential equations, dynamical systems, and control science, 867–890, Lecture Notes in Pure and Appl. Math., 152, Dekker, New York, 1994.

[15] G. Tessitore. Some remarks on the Riccati equation arising in an optimal control problem with state- and control-dependent noise. SIAM J. Control Optim. 30 (1992), no. 3, 717–744.

[16] G. Tessitore. Existence, uniqueness and space regularity of the adapted solutions of a backward SPDE. Stochastic Anal. Appl. 14 (1996), no 4, 461–486.

[17] J. M. A. M. van Neerven, M. C. Veraar, L. Weis. Stochastic integration in UMD Banach spaces. Ann. Probab. 35 (2007), no. 4, 1438–1478.

[18] J. Yong, X.Y. Zhou. Stochastic controls. Hamiltonian systems and HJB equations. Applications of Mathematics (New York), 43. Springer, New York, 1999.

[19] X.Y. Zhou. On the necessary conditions of optimal controls for stochastic partial differential equations. SIAM J. Control Optim. 31 (1993), no. 6, 1462–1478.